ZA-Information 27

November 1990

Dieses Dokument wurde von einer gedruckten Vorlage eingescannt und mit einer OCR-Software weitgehend automatisch bearbeitet. Kleinere Übertragungsfehler sind daher möglich.

ZA-Information 27

November 1990

Herausgeber:
Zentralarchiv für empirische Sozialforschung
Universität zu Köln
Das Zentralarchiv ist Mitglied der GESIS

Direktor: Prof. Dr. E. K. Scheuch
Geschäftsführer: E. Mochmann

Anschrift:
Bachemer Straße 40
5000 Köln 41

Telefon:
Zentrale 0221 / 4 76 94 - 0
Telefax - 44
Redaktion - 50

EARN/BITNET: ZA at DK0ZA1

Redaktion:
Franz Bauske

ISSN: 0723-5607
© Zentralarchiv

Die ZA-INFORMATION erscheint jeweils im Mai und November eines Jahres.
Sie wird kostenlos an Interessenten und Benutzer des Zentralarchivs abgegeben.

inhalt

ZA-Information 27 **November 1990**

Mitteilungen der Redaktion 4

Berichte aus dem Archiv

Neuauflage des Datenbestandskatalogs des Zentralarchivs 5

Das Zentralarchiv auf dem Soziologentag in Frankfurt 16

ALLBUS: Eine Übersicht zu dem gegenwärtigen Datenangebot 18

Das Thema Aids in den Datensätzen des Zentralarchivs 30

Studien aus der ehemaligen DDR im Zentralarchiv 44

Archivexperten diskutieren Datenbankprobleme 47

Überlegungen zur Konzeption des Daten- und Informationsmanagements für den ALLBUS von Michael *Terwey* 50

Forschungsnotizen

Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen
von Wolfgang ***Ludwig-Mayerhofer*** 62

Lassen sich mit SPSSx-Matrix anwenderspezifische Analyseprobleme lösen?
Ein Anwendungstest am Beispiel der multinomialen logistischen Regression
von Steffen M. ***Kühnel*** 89

Die komplementäre Anwendung zweier Verfahren:
Korrespondenzanalyse und Konfigurationsfrequenzanalyse
von Jörg ***Blasius*** und Erwin ***Lautsch*** 110

Nun kann auch die soziologische Methodik zusammenwachsen
von Michael ***Hader*** 134

Eine Wiederbesichtigung des Forschungsbetriebs "Soziologie" in der UdSSR
von Wolfgang ***Teckenberg*** 141

Berichte, Ankündigungen und Mitteilungen

Conference on Correspondence Analysis 154

Empirische Sozialforschung über Entwicklungsländer (ESE) 155

Bücherhinweise 156

Die Ost-Stichprobe des SOEP:
Konzept und Durchführung der "Basiserhebung 1990" in der DDR 163

Drogenkonsum und Drogenpolitik: Europäische Zusammenarbeit im Bereich der
sozialwissenschaftlichen Drogenforschung 167

7th European Meeting of the Psychometric Society in Trier 168

Integrating the European Data Base 169

Auf dem Weg zum Europäischen Sozialwissenschaften-Netz (ESONET) 170

Multivariate Analyse kategorialer Daten:
Friihjahrsseminar des Zentralarchivs vom 18.2. bis 8.3.1991 172

Bei Beiträgen, die nicht von Mitarbeitern des Zentralarchivs verfaßt wurden, wird die Anschrift des Autors im Anschluß an den jeweiligen Artikel angegeben.
Die Inhalte der Beiträge entsprechen der Meinung der Autoren und geben nicht unbedingt in jedem Punkt die Ansicht der Redaktion wider.

Mitteilungen der Redaktion

Der erweiterten Neuauflage des Datenbestandskatalogs sind die ersten Seiten dieser Ausgabe gewidmet. Das Nachschlagewerk über die Bestände des Zentralarchivs stellen wir auf den folgenden Seiten ausführlich vor. Der neue Katalog enthält eine detaillierte und aktuelle Übersicht über alle Datensätze und ist besonders im Bereich der Register durch umfangreiche Ergänzungen und Neuerungen gekennzeichnet. Unseren Lesern können wir das Buch zu Sonderkonditionen anbieten, ebenso wie die Veröffentlichung von ***Blasius und Dangschat*** zum Thema Gentrification. Vorgestellt wird auch die neue Ausgabe der *"Empirischen Sozialforschung"*, die über die quantitative Forschung des letzten Jahres informiert.

Erneut richten wir den Blick nach Osten. Im Zentralarchiv sind erste Datensätze aus der ehemaligen DDR eingetroffen. Michael ***Hader*** berichtet über die Vergangenheit und die Perspektiven der empirischen Sozialforschung in der DDR. Wolfgang ***Teckenberg*** beschreibt die Lage der Soziologie in der Sowjetunion und knüpft damit an seinen früheren Beitrag in dieser Zeitschrift an.

Angeregt durch den Beitrag von Steffen ***Kühnel et al.*** in der ZA-Information 25 und den Folgebeitrag von Dieter ***Urban*** in der letzten Ausgabe hat uns Wolfgang ***Ludwig-Mayerhofer*** ebenfalls einen Artikel zur logistischen Regression eingereicht. Er präsentiert ein sehr anschauliches Beispiel zur Analyse einer **ordinalen** abhängigen Variablen mit einem logistischen Modell.

Ein Problem solcher Analysemethoden ist sicherlich in ihrer geringen Bekanntheit und Verfügbarkeit zu sehen. Dem will u.a. auch das Frühjahrsseminar 1991 begegnen, in dem H. J. ***Andress*** aus Bielefeld u. a. in die Analyseverfahren mit loglinearen Modellen einführt. Schließlich geht Steffen ***Kühnel*** in seinem Beitrag noch der Frage nach, ob die SPSS^x-Prozedur "MATRIX" geeignet ist, derartige Analysemodelle selbst zu programmieren.

Franz ***Bauske***

Neuauflage des Datenbestandskatalogs des Zentralarchivs

"Daten der empirischen Sozialforschung"

Der neue Datenbestandskatalog des Zentralarchivs mit Beschreibungen von Daten der empirischen Sozialforschung von 1945 bis 1990 und Daten der historischen Sozialforschung ist zur Zeit im Druck.

Der Datenbestandskatalog bietet eine vollständige Übersicht über die archivierten maschinenlesbaren Datensätze, die bis Mitte des Jahres 1990 akquiriert und aufbereitet worden sind. Gegenüber der ersten Auflage wurden vor allem ein Namensregister der Primärforscher, eine Auflistung der Erhebungsinstitute sowie die inzwischen erfolgten Neuzugänge an empirischen Studien hinzugefügt. Darüber hinaus sind das Datenangebot des Zentrums für Historische Sozialforschung (ZHSF) und des Entwicklungsländerprojektes in den Datenbestandskatalog aufgenommen worden.

Ziel des Datenbestandskatalogs ist es, einen Gesamtüberblick über die ausleihbaren Datensätze zu vermitteln und Interessenten die Möglichkeit einer ersten Auswahl zu bieten, ohne daß eine direkte Suche in den Beständen des Archivs notwendig ist. Alle vorgestellten Archivbestände sind auf maschinenlesbaren Datenträgern gespeichert. Sie sind der wissenschaftlichen Öffentlichkeit zugänglich, wenn auch zum Teil mit Auflagen von Seiten der Datengeber versehen. Die Weitergabe geschieht entsprechend den jeweils individuellen Anforderungen und datenverarbeitungstechnischen Spezifikationen des Ausleihers (Magnetband oder Diskette).

Der Datenbestandskatalog wird vom Campus Verlag in der Reihe des Zentralarchivs " Beiträge zur Empirischen Sozialforschung" verlegt. Im Frühjahr 1991 wird er im Buchhandel erscheinen. Unseren Nutzern können wir den über 800 seitigen Katalog (DIN A 4 Format) zum Sonderpreis von 48.- DM anbieten. Siehe Bestellformular am Ende dieses Beitrags.

Auszug aus dem Inhaltsverzeichnis des Datenbestandskatalogs:

Das Zentralarchiv und sein Dienstleistungsangebot für die Sozialforschung
Benutzungsordnung
Gebührenordnung
Die Datenbestände des Zentralarchivs
Erläuterungen zu den Studienbeschreibungen
Register
- Register der Studientitel
- nach Inhaltskategorien (ZA-Kategorien) sortiert
- alphabetisch sortiert
- numerisch sortiert
- Erhebungszeitraum
- Geographisches Register
- Primärforscher
- Erhebungsinstitute und Primärforscherinstitute
- Stichwortverzeichnis
Hinweise für Datengeber
Hinweise zur Auswahl und Bestellung von Daten

Die Beschreibung der Datenbestände des Zentralarchivs

Im Mittelpunkt der Veröffentlichung steht die inhaltliche und methodische Beschreibung der archivierten Untersuchungen. Alle im Zentralarchiv archivierten Studien werden in Form von Studienbeschreibungen vorgestellt. In diesen Darstellungen wird ein relativ kurzer, aber weitgehend vollständiger Überblick über den Inhalt einer Untersuchung gegeben. Jede Studienbeschreibung ist nach einem standardisierten Gliederungsschema aufgebaut, das auf der folgenden Seite in einem Beispiel dargestellt wird. Die Beschreibungen sind nach Archivnummern aufsteigend abgedruckt, wodurch der Zugriff aufgrund von Registerverweisen leicht möglich ist.

ZA-Nr. **1713** Zugang: A

Eurobarometer 28 (Relations with Third World Countries and Energy Problems)

Erhebungszeitraum
Oktober 1987 bis November 1987

Primärforscher
K. Reif,
Kommission der Europäischen Gemeinschaften, Brüssel

Datenerhebung
DIMARSO, Brüssel;
EMNID, Bielefeld;
Gallup Markedsanalyse, Hellerup;
Institut de Sondages Lavialle, Issy-Les-Moulineaux;
ICAP Hellas, Athen;
Irish Marketing Surveys, Dublin;
DOXA, Mailand;
ILRES, Luxemburg;
Instituto de Investigacion Gallup, Madrid;
NIPO, Amsterdam;
NORMA, Lissabon;
Social Surveys (Gallup Poll), London

Inhalt
Die Befragung hat 5 Schwerpunkte:
1.) Allgemeine Zufriedenheit
2.) Fragen zur Entwicklungshilfe
3.) Beurteilung von Energiefragen
4.) Einstellung zur Europäischen Vereinigung
5.) Einstellung zur Unterstützung von Sportveranstaltungen durch die EG
6.) Beurteilung der europäisch-amerikanischen Beziehungen
7.) Fragen zur europäischen Jugend

Themen: 1.) Wahlberechtigung am Wohnort; allgemeine Zukunftsaussichten für das kommende Jahr; erwartete Zunahme von Streiks und internationalen Konflikten; Kriegsfurcht (Skalometer); Beurteilung der wirtschaftlichen Lage der Bundesrepublik und der eigenen finanziellen Situation; allgemeine Lebenszufriedenheit; Demokratiezufriedenheit; eigene Meinungsführerschaft; Häufigkeit von politischen Diskussionen; Postmaterialismus; Kenntnis des europäischen Krebsbekämpfungsprogramms; wichtigste Interessengebiete; Bereiche persönlichen Engagements; Erziehungsziele; wichtigste Jugendprobleme; Mitgliedschaften; Benutzung von Begegnungsstätten und Jugendzentren in der Gemeinde; Selbsteinschätzung der Sprachkenntnisse; Fremdsprachenkenntnisse und Lehrinstitution; Sprache zu Hause; Einschätzung der Nützlichkeit und Wichtigkeit des Erlernens von Fremdsprachen; Beurteilung des Sprachunterrichts in Schulen; Beschreibung des Haushaltstyps und Grund für die Familiensituation; erhaltene finanzielle Unterstützungsleistungen beim Aufziehen von Kindern.
2.) Einstufung der Wichtigkeit ausgewählter politischer Probleme und Präferenz für nationale oder europäische Problemlösung (Skala); Einschätzung der zukünftigen Entwicklung von internationalen Spannungen, Lebensstandard, Hunger, technischem Fortschritt im Dienste der ärmeren Länder und des Verhältnisses der Industrienationen zu anderen Ländern; rücksichtsvolle Nutzung von Ressourcen der Erde; Kenntnis von Staaten, in denen Menschen an Hunger sterben; Einschätzung der Fortschritte in ausgewählten Entwicklungsländern und Angabe von Ländern, die für nicht entwicklungsfähig gehalten werden; Beurteilung der Medienberichterstattung über die Dritte Welt; Bereiche, über die man sich mehr Informationen wünscht; allgemeine Beurteilung der Verbesserung bzw. Verschlechterung der Lage von Entwicklungsländern; Einstellung zur Entwicklungshilfe; Hauptziel der Beziehungen zu Ländern der Dritten Welt; Beurteilung der Erfolgschancen von Entwicklungspolitik; wichtigste Gründe für Entwicklungspolitik; eigene Erfahrungen in Entwicklungsländern; Einschätzung des Einflusses der wirtschaftlichen Entwicklung in Ländern der Dritten Welt auf das Leben im eigenen Lande; Einschätzung der politischen und wirtschaftlichen Abhängigkeiten zwischen Industrieländern und Ländern der Dritten Welt (Skala); Kenntnisse über Entwicklungshilfeleistungen durch die Landesregierung, die Europäische Gemeinschaft, die UNO, durch Unternehmen und Industrie sowie nicht-staatliche Hilfsorganisationen; Institutionen, die die effektivste Entwicklungshilfe leisten; Hilfsaktionen für die Dritte Welt in der Nachbarschaft bzw. am Wohnort; erhaltene Anfragen nach Geldspenden; Teilnahme an Dritte Welt-Veranstaltungen; Mitgliedschaft in einer Dritte Welt-Gruppe; Bereitschaft, Zeit für Aktionen aufzuwenden und Geldspenden für die Dritte Welt zu geben.
3.) Einschätzung des derzeitigen sowie zukünftigen Energieproblems des Landes; Kenntnis des Selbstversorgungsgrads sowie Einschätzung der zuverlässigsten Energielieferanten des Landes; präferierte Politik zur Sicherung des zukünftigen Energiebedarfs; Einschätzung der Energieverschwendung im eigenen Land; Einstellung zur staatlichen Reglementierung bei der Verringerung des Energieverbrauchs; Einstellung zur staatlichen Subventionierung der Erforschung neuer Energieformen bzw. zur Subventionierung von energiesparenden Einrichtungen; preisstabilste, zuverlässigste und umweltverträglichste Energieform der Zukunft; Einschätzung des Risikos ausgewählter Industrieanlagen; Einstellung zur Weiterentwicklung von Atomkraftwerken; Ausbau der Atomkraft oder Einschränkung des Energieverbrauchs; Einschätzung der Ernsthaftigkeit des Energieproblems; wichtigste Gefahren durch ein Atomkraftwerk und den radioaktiven Abfall; erwartete Entwicklung des Stromverbrauchs in den nächsten Jahren; Einstellung zur Verringerung des Erdölanteils bei der Elektrizitätsgewinnung; Präferenz für Kohle oder Kernenergie zur Deckung des zusätzlichen Energiebedarfs; eigenes Energiesparverhalten und Bereiche, in denen noch zusätzliche Einsparmöglichkeiten gesehen werden; tatsächlich verwendete und gewünschte Energieart beim Heizen und Kochen im Haushalt; Besitz langlebiger Wirtschaftsgüter; Einstellung zur Atomkraft (Skala); Kenntnis internationaler Schutzmaßnahmen im Falle eines schweren Atomunfalls; schwerwiegendste Folgen einer angenommenen Ölverknappung im Lande.
4.) Kenntnis und Ansehen der EG-Kommission; allgemeine Einstellung zur Vereinigung Westeuropas; tatsächlich festgestellte und gewünschte Entwicklung der Vereinigung Europas; Einstellung zur Mitgliedschaft des Landes in der EG; Einschätzung der Vorteilhaftigkeit des freien Verkehrs von Personen, Waren und Gütern innerhalb der Gemeinschaft bei Verwirklichung des Europäischen Marktes 1992 (Skala); allgemeine Befürwortung des Binnenmarkts und einer weiteren EG-Einigung; Angabe der Richtung, in die die Einigung fortschreiten sollte; Einschätzung der Vorteilhaftigkeit der EG-Mitgliedschaft des Landes; Bedauern eines Scheiterns der Europäischen Gemeinschaft; Beurteilung des Einvernehmens und zukünftige Rolle des Parlaments; Wahlbeteiligungsabsicht zwischen den Staaten der EG; Bekanntheit des Europäischen Parlaments und allgemeiner Eindruck vom Parlament; gegenwärtige bei einer Wahl zum Europäischen Parlament; Einstellung zur Stärkung der Befugnisse des Europäischen Parlaments; Einstellung zur Gründung einer europäischen Regierung und Ressorts, für die eine europäische Regierung verantwortlich sein sollte; Einstellung zu Mehrheitsbeschlüssen in der EG und zu einem Verfassungsentwurf für ein vereintes Europa.
5.) Kenntnis von EG-unterstützten Sportveranstaltungen; Einstellung zur finanziellen Unterstützung von Sportveranstaltungen durch die EG; Beitrag der Sportveranstaltungen zur Erhöhung des Bekanntheitsgrads der EG.
6.) Einstellung zu den USA; Beurteilung der Bedeutung der europäisch-amerikanischen Bindung; vermutete Einstellung der USA zu einem vereinigten Westeuropa; vermutete Entwicklung der europäisch-amerikanischen Beziehungen in den nächsten Jahren.
7.) An Befragte im Alter von 15 bis 34 Jahren wurden folgende Fragen zusätzlich gestellt: Interesse an Arbeit oder Ausbildung in einem europäischen sowie außereuropäischen Land; wichtigste Gründe für einen Auslandsaufenthalt; Wichtigkeit staatlicher Unterstützung des Jugendaustauschs; erwartete Zunahme der Mobilität durch den Europäischen Markt; Veränderungen der Berufsaussichten durch die zunehmende Konkurrenz im eigenen Berufsfeld; Umzugsabsicht in den nächsten 10 Jahren; mögliche Gründe für einen Berufswechsel in den nächsten Jahren; Arbeitslosigkeit und Mobilität; Wohnen in der Gegend, in der man aufgewachsen ist; Lesen ausländischer Zeitschriften; Auslandsaufenthalte; ähnliche Berufstätigkeit wie die Eltern; Urlaubsaufenthalt an anderen Orten; Großunternehmen mit besonders europäischer Gesinnung.
8.) Selbsteinschätzung auf einem Links-Rechts-Kontinuum; Einstellung zu Veränderungen der Gesellschaftsordnung; Parteiverbundenheit; Wahlverhalten bei der letzten Wahl; Parteipräferenz; Ortsgröße.
In Irland wurde zusätzlich gefragt: Einstellung zur irischen Sprache.
In Belgien wurde zusätzlich gefragt: Sprache im Haushalt.
Indizes: Kognitive und politische Mobilität; Postmaterialismus-Index; politische EG-Unterstützung; Links-Rechts-Index.

Grundgesamtheit und Auswahl
Untersuchungsgebiet: Belgien, Bundesrepublik Deutschland, Dänemark, Frankreich, Griechenland, Irland, Italien, Luxemburg, Niederlande, Portugal, Spanien, Vereinigtes Königreich.
Verschiedene Auswahlverfahren (Quotenauswahl und mehrstufige Zufallsauswahl) je nach Land. Auswahl von Personen im Alter von 15 Jahren und älter.

Register zum Datenbestandskatalog

Der Schlußteil des Datenbestandskatalogs enthält auf über 100 Seiten mehrere Register, die nach verschiedenen Kriterien erstellt worden sind. Damit wird es dem Benutzer möglich, eine auf seine individuelle Fragestellung abgestimmte Suchstrategie anzuwenden und geeignete Datensätze schnell zu identifizieren.

Register der Studientitel

Ein erster Überblick über das Themenspektrum der im Zentralarchiv erhälüichen Datensätze kann anhand der Studientitel gewonnen werden. Das Register enthält die Titel aller archivierten Studien in drei verschiedenen Sortierweisen:

1. Sortierung entsprechend dem **inhaltlichen** Schwerpunkt der Studie. Sortierkriterium ist dabei die **ZA-Kategorie**, die eine Verschlüsselung des Hauptinhalts der jeweiligen Untersuchung darstellt,
2. **alphabetische** Sortierung und
3. **numerische** Sortierung, nach Archivnummern aufsteigend geordnet.

(zu diesen drei Punkten finden sich anschauliche Beispiele aus dem Datenbestandskatalog auf den nächsten drei Seiten)

Studientitel nach ZA-Kategorien sortiert

ZA-Kat 10 Staat
0025 Einstellung zur Monarchie (Niederlande)
0326 Politik und Demokratie in der Bundesrepublik (Studenten, Januar 1968)
0521 Politik und Demokratie in der Bundesrepublik (Bevölkerung, 1968)
0522 Politik und Demokratie in der Bundesrepublik (Jugendliche, 1968)
0523 Politik und Demokratie in der Bundesrepublik (Studenten, Juni 1968)
0528 Jugend und Demokratie 1968
0970 Ausgewählte politische Fragen (Deutschlandbild) in acht Ländern (Dänemark, Frankreich, Griechenland, Großbritannien, Italien, Niederlande, Schweden, USA)
1253 Jugend und Staat
1486 Ansprüche der Bürger an den Staat (1985/86)
1487 Ansprüche der Bürger an den Staat (Integrierte Studie 1985/86 und 1989)
1540 Public Attitudes Toward Democracy in Post-War Germany

ZA-Kat 11 Staatliche Institutionen
0480 Verwaltung und Öffentlichkeit (Bevölkerung)
0481 Verwaltung und Öffentlichkeit (Meinungsführer)
0837 Jungparlamentarier in Bonn
0883 Bundesratsfunktionen
0887 Einstellung zu Beamten
0920 Sozioökonomische, politische und fiskalische Daten für die Bundesländer (1960 und 1970)
0962 Kommunikationsanalyse Nürnberg
0965 Informationssysteme und Informationsverhalten (Beobachtungsstudie und Klientenbefragung)
1015 Steuerzahler und Finanzamt (Steuerzahlerbefragung)
1016 Steuerzahler und Finanzamt (Sachbearbeiterbefragung)
1034 Informationssysteme und Informationsverhalten (Bevölkerung)
1035 Informationssysteme und Informationsverhalten (Sachbearbeiter)
1345 Staatliches und kommunales Archivwesen der Bundesrepublik Deutschland
1583 Einstellung zur Volkszählung (März-April 1987)
1584 Einstellung zur Volkszählung (April-Mai 1987)
1585 Einstellung zur Volkszählung (Mai 1987)
1586 Einstellung zur Volkszählung (Juni-Juli 1987)
1587 Einstellung zur Volkszählung (November-Dezember 1987)
1588 Einstellung zur Volkszählung (Panel: 1. Welle, April-Mai 1987)
1589 Einstellung zur Volkszählung (Panel: 2. Welle, Juni-Juli 1987)
1590 Einstellung zur Volkszählung (Panel: 3. Welle, November 1987-Januar 1988)

ZA-Kat 12 Internat Institutionen, Beziehungen, Verhältnisse
0069 Internationale Beziehungen (Großbritannien)
0070 Internationale Beziehungen (Frankreich)
0071 Internationale Beziehungen (Deutschland)
0072 Internationale Beziehungen (Italien)
0073 Europa in der öffentlichen Meinung 1962 (Frankreich)
0074 Europa in der öffentlichen Meinung 1962 (Belgien)
0075 Europa in der öffentlichen Meinung 1962 (Luxemburg)
0076 Europa in der öffentlichen Meinung 1962 (Niederlande)
0077 Europa in der öffentlichen Meinung 1962 (Deutschland)
0078 Europa in der öffentlichen Meinung 1962 (Datenpool)
0284 Einstellung zur fortschreitenden europäischen Integration
0333 Politische Fragen (Europäische Vereinigung)
0380 Internationale Beziehungen (Juni 1961, Form A)
0381 Internationale Beziehungen (Juni 1961, Form B)
0421 Entwicklungshilfe
0428 Internationale Beziehungen (Februar 1964)

Studientitel in alphabetischer Abfolge

1286 A Survey of Europe Today (Belgien)
1295 A Survey of Europe Today (Dänemark)
1290 A Survey of Europe Today (Deutschland)
1296 A Survey of Europe Today (Finnland)
1285 A Survey of Europe Today (Frankreich)
1293 A Survey of Europe Today (Großbritannien)
1294 A Survey of Europe Today (Irland)
1289 A Survey of Europe Today (Italien)
1288 A Survey of Europe Today (Luxemburg)
1287 A Survey of Europe Today (Niederlande)
1298 A Survey of Europe Today (Norwegen)
1292 A Survey of Europe Today (Österreich)
1300 A Survey of Europe Today (Portugal)
1297 A Survey of Europe Today (Schweden)
1291 A Survey of Europe Today (Schweiz)
1299 A Survey of Europe Today (Spanien)
0356 Abendgymnasiasten in Köln
8004 Abgeordnete der Assemblee Nationale Constituante 1848-1849
8003 Abgeordnete der Frankfurter Nationalversammlung 1848-1849
8007 Abgeordnete der Reichstage der Weimarer Republik 1919-1933
8006 Abgeordnete der Reichstage des Kaiserreichs 1867/71-1918
0187 Abgrenzung und Konstanz des Mittelstandes
8030 Abitur 1917 bis 1971. Inhaltsanalyse von Abituraufsätzen
1400 Abiturienten in Schleswig-Holstein 1981
1459 Absolventenbefragung 1983
0872 Abweichendes Verhalten in der Schule (Schülerbefragung)
8043 Adelsstruktur und Reichspolitik. Untersuchungen zur politischen und gesellschaftlichen Relevanz kaiserlicher Standeserhebungen 1470-1740
0375 Ägyptische Studenten
1484 Ältere Augsburger (Heimbewohner)
1483 Ältere Augsburger (Privathaushalte)
0419 Ärzte
0950 Aggregatdaten für die Wahlkreise der Bundesrepublik (1949-1969) in der Einteilung der Bundestagswahl 1965
0951 Aggregatdaten für die Wahlkreise der Bundesrepublik (1972-1976) in der Einteilung der Bundestagswahl 1972
1184 Aggregatdaten für die Wahlkreise der Bundesrepublik (1976-1980)
1159 Aggregate Dataset Eastern Europe
1758 AIDS im öffentlichen Bewußtsein der Bundesrepublik (1987)
1757 AIDS im öffentlichen Bewußtsein der Bundesrepublik (1988)
0423 Akademiker in Afghanistan
1041 Akademiker in Deutschland
1728 Aktionsgrundlagen 1987 der Bundeszentrale für gesundheitliche Aufklärung (Effizienzkontrolle)
1183 Aktionsräume von Stadtbewohnern
1052 Alkohol und Fahren
1163 Alkoholische Getränke
1369 Alkoholismus als Karriere
0406 All-India Political Poll
1000 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1980
1160 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1982
1340 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1984
1601 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1984 (Test-Retest-Studie)
1500 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1986
1669 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1986 (Non-Response-Studie)
1670 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1988
1335 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS Kumulierter Datensatz 1980-1986
1795 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS Kumulierter Datensatz 1980-1988
0037 Allgemeiner und politischer Traditionalismus
0573 Alte Menschen im Landkreis Leer (Altenwohnungen und Altenheime)
0572 Alte Menschen im Landkreis Leer (Privathaushalte)
0158 Alte Menschen in Düsseldorf-Mettmann (Altenheimbewohner)
0157 Alte Menschen in Düsseldorf-Mettmann (Privathaushalte)
1261 Altenuntersuchung (Dreieich)
1262 Altenuntersuchung (Kreis Offenbach)
0146 Altersbild und Altersvorsorge der Arbeiter und Angestellten
1182 Altruistisches Verhalten von freiberuflich Tätigen

Studientitel in numerisch aufsteigender Folge

1760 Wahrnehmung der Reaktorkatastrophe von Tschernobyl durch die deutsche Bevölkerung (Mai 1987)
1761 Wahrnehmung der Reaktorkatastrophe von Tschernobyl durch die deutsche Bevölkerung (Mai 1988)
1762 Politbarometer 1988 (Kumulierter Datensau)
1763 Einstellungen zu aktuellen Fragen der Innenpolitik (1989)
1764 Gorbatschow-Besuch
1765 Europawahl 1989
1766 Wahl zum Abgeordnetenhaus in Berlin 1989
1767 Kommunalwahl in Frankfurt 1989
1768 Politbarometer (Januar 1989)
1769 Politbarometer (Februar 1989)
1770 Politbarometer (März 1989)
1771 Politbarometer (April 1989)
1772 Politbarometer (Mai 1989)
1773 Politbarometer (Juni 1989)
1780 Berliner SPD-Mitgliederbefragung 1988
1782 Media-Analyse (MA 89, Funkmedien)
1783 Media-Analyse (MA 89, Pressemedien)
1785 Meinungsbildner 1989
1786 Ausländer in Deutschland 1987
1787 Ausländer in Deutschland 1989
1788 Wohnwelten in Deutschland 1987
1789 Das Leben zu zweit - Sexualverhalten 1988
1790 Qualifikation und Berufsverlauf (1985/86)
1792 Soll und Haben 2
1793 Einstellung zur Technik (1989)
1794 Die Funktion und Rolle des Berufsberaters
1795 Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS Kumulierter Datensatz 1980-1988
1796 Sicherheitspolitische Einstellungen in der Bundesrepublik 1980
1797 Sicherheitspolitische Einstellungen in der Bundesrepublik 1981
1798 Sicherheitspolitische Einstellungen in der Bundesrepublik 1984
1799 Sicherheitspolitische Einstellungen in der Bundesrepublik 1985
1804 Frauen in Beruf, Haushalt und Öffentlichkeit (Schleswig-Holstein-Survey 1989)
1805 Perspektiven des ländlichen Raumes
1806 Meinungen der sowjetischen Bürger über die Deutschen 1989
1807 Frühjahrsstudie 1984
1808 Herbststudie 1984
1809 Herbststudie 1985
1810 Frühjahrsstudie 1986
1811 Herbststudie 1986
1812 Arbeitsdatensatz: Statistische Datenanalyse mit Mikrocomputern
1813 Verhinderung und Bekämpfung von Gewalt (Januar-Februar 1989)
1814 Verhinderung und Bekämpfung von Gewalt (März-April 1989)
1815 Rangliste der westdeutschen Hochschulen
1817 Wahlverhalten, religiös-konfessionelle Einstellungen und Einstellungen zur wirtschaftlichen Lage
1818 Einstellungen zu Sexualität und AIDS
1819 International Environmental Survey 1980 (Germany)
1826 Typologie der Wünsche 1989
1827 Regionaldaten VZ 1961 (Bundesländer)
1828 Regionaldaten VZ 1961 (Regierungsbezirke)
1829 Regionaldaten VZ 1961 (Kreise)
1830 Einstellungen zu AIDS (Ärztebefragung)
1832 Regionaldaten VZ 1950 (Bundesländer)
1833 Regionaldaten VZ 1950 (Regierungsbezirke)
2000 Historische Indikatoren der westeuropäischen Demokratien (HIWED)
5003 The Role of Women in the Informal Sector: The Case of Ajegunle, Nigeria
5004 Die sozioökonomische Rolle der Frau als Kleinunternehmerin im informellen Sektor in Abidjan, Elfenbeinküste
5005 Farm Survey and Small-Scale Irrigation Project in Gwembe Valley, Zambia
5006 Strukturen und Funktionen des informellen Sektors in Ghana
5007 Sozioökonomische Situation von Kleinbauern in Nigeria
5008 Small-Scale Industries and Economic Development in Ghana: Business Behavior and Strategies in Informal Sector Economies
5009 Process of Differentiation in Small-Scale Agricultural Family Enterprises in Pelarco, Chile (1st Wave, 1984)
5010 Die chilenische Küstenkordillere: Ländliche Armut und Ressourcendegradation
5011 Konfliktlinien der ägyptischen Gesellschaft 1986-1988: Remigration, Segregation, Islamismus (Haushaltszensus in 6 ägyptischen Dörfern: Haushaltsdaten)
5012 Konfliktlinien der ägyptischen Gesellschaft 1986-1988: Remigration, Segregation, Islamismus (Haushaltszensus in 6 ägyptischen Dörfern: Migrantendaten)
5013 Konfliktlinien der ägyptischen Gesellschaft 1986-1988: Remigration, Segregation, Islamismus (Befragung von zurückgekehrten Migranten in 6 ägyptischen Dörfern)
5014 Der Einfluß des Fernsehens in ländlichen Gemeinden Tunesiens (Panel: 1. Welle, 1982)
5015 Der Einfluß des Fernsehens in ländlichen Gemeinden Tunesiens (Panel: 2. Welle, 1983)
5016 Der Einfluß des Fernsehens in ländlichen Gemeinden Tunesiens (Panel: 3. Welle, 1985)
5018 Small-Scale Manufacturing and Repair Activities in the Urban Area of Techiman, Ghana
5023 Konfliktlinien der ägyptischen Gesellschaft 1986-1988: Remigration, Segregation, Islamismus (Frauen von zurückgekehrten Arbeitsmigranten)
5026 Process of Differentiation in Small-Scale Agricultural Family Enterprises in Pelarco, Chile (2nd Wave, 1988) and Supplementary in Chilean Municipalities
8001 Soziologie des Widerstandes gegen den Nationalsozialismus: Das Beispiel Düsseldorf 1933 bis 1945
8002 Die Bevölkerung der Exulantenstadt Glückstadt 1616-1652
8003 Abgeordnete der Frankfurter Nationalversammlung 1848-1849
8004 Abgeordnete der Assemblee Nationale Constituante 1848-1849
8005 Ökologische Daten zu Frankreich um 1848
8006 Abgeordnete der Reichstage des Kaiserreichs 1867/71-1918
8007 Abgeordnete der Reichstage der Weimarer Republik 1919-1933
8008 Datenbank Abgeordnete deutscher Nationalparlamente 1848-1933
8009 Sozialdemokratischer Parteitag in Hamburg 1897
8010 Sozialdemokratische Reichstagskandidaten und Reichstagswahlen von 1898 bis 1918 (BIOKAND)
8011 Sozialdemokratische Parlamentarier in den deutschen Reichs- und Landtagen 1867 bis 1933 (BIOSOP)
8012 Die Hochschullehrer der Technischen Hochschule Berlin von 1879 bis 1945
8013 Wahl- und Sozialdaten der Kreise und Gemeinden des Deutschen Reiches von 1920 bis 1933
8014 Getreideumsatz, Getreide- und Brotpreise in Köln vom 14. bis 18. Jahrhundert
8015 Handelspolitische Petitionen an die Frankfurter Nationalversammlung 1848/49
8016 Textilarbeiterschaft in der Industrialisierung
8017 Die Schiffahrts- und Handelsbeziehungen Schleswig-Holsteins nach Lateinamerika 1815-1848
8018 Zum Entstehen einer politischen Öffentlichkeit in der Schweiz. Die Mitglieder und Gäste der helvetischen Gesellschaft
8019 Untersuchungen zum Teilbuch (Steuerbuch) der Stadt Bern von 1389
8020 Untersuchungen zum Tellbuch (Steuerbuch) der Stadt Bern von 1494
8021 Familie und soziale Plazierung in den Bielefelder Unterschichten im 19. Jahrhundert
8022 Strukturwandel der Familie in Österreich seit dem 17. Jahrhundert
8023 Die Struktur des niedersächsischen Landvolkes im ausgehenden 17. Jahrhundert
8024 Die Bedeutung des Systems der zentralen Orte für den sozialen und wirtschaftlichen Wandel im Herzogtum Schleswig in der Frühindustrialisierung
8025 Die räumliche Sozialstruktur der Stadt Bern 1764 und 1896
8026 Schweizer Nationalwahlen 1848-1919
8027 Arbeitsvertragsbrüche als Verletzung der Arbeitspflicht im "Dritten Reich"

Primärforscher-Register

Das Primärforscher-Register enthält die Namen aller Personen, die hauptverantwortlich an Konzeption und Durchführung der Untersuchungen beteiligt waren. Die Zahlen verweisen auf die entsprechenden ZA-Studien-Nummern.

Aarebrot, F. H. 1411 1412 1413 1414
Abrams, M. 0760 0765 1431
Abshagen, R. 0254
Adick,C. 1624
Adrian, W. 0519 0520
Aich, P. 0195
Aitkin, D. 1401
Alber, J. 2000
Albrecht, R. 5011 5012 5013 5023
Aleksic, P. 0271 0625
Allerbeck, K. R. 0363 0757 0765 1040 1188 1189 1190 1191 1335 1431 1436 1795
Almond, G. A. 0027 0028 0029 0030 0031 0624
Alpheis, H. 1580
Anderseck, K. 0864
Andreenkov, V. G. 1666
Andrey, G. 8026
Anheier, H. K. 1671 5003 5008
Aryeetey, E. 5018
Asher, H. A. 1401 1402 1403 1404 1405 1406 1407 1408
Badura, B. 0837
Bächler, H. M. 8034
Bäthge, M. 0878
Baker, K. L. 0150
Baldermann, B. 0952
Banhegyi, F. 0023
Bargel, T. 1096 1097
Barnes, S. H. 0761 0765 1188 1189 1431
Bartholomeyczik, E. 0958 1022 1023 1024
Bartholomeyczik, S. 1022 1023 1024
Bartscher, U. 0490
Batori, I. 8042
Bauer, H. 8054
Bauer, M. 0630
Bauks, F. W. 1512
Baumert, G. 0055 0056 0057 0058 0059 0060 0061 0062 0084 1575 1576 1577
Baunach, N. 0357
Baur, R. 0893 1151
Bechmann, G. 1225 1226
Beck, G. 0529
Becker, E. 0156
Becker, H. 1389
Beckers, H. J. 0280
Beckmann, M. 1022 1023 1024
Beine, H. 1344
Bellers, J. 1344
Benad, G. 1365
Bentler, U. 1663
Berger, M. 0519 0520 0635 0636 0637 0823 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1244 1245 1246 1247 1248 1249 1250 1251 1275 1276 1277 1278 1279 1280 1281 1371 1372 1403 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1551 1552 1554 1555 1556 1581 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1677 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1762 1763
Bernhard, W. 1022 1023 1024
Best, H. 8003 8004 8005 8006 8007 8008 8015
Bick, W. 0965 1034 1035 1186
Bimmer, B. 0875
Binder-Krauthoff, K. 0492
Bingemer, K. 0532
Blankenburg, E. 0037
Blasius, J. 1625 1740 1741 1742 1827 1828 1829 1832 1833
Blau, P. 0822
Blücher, V. Graf 0772 0925
Blume, O. 0198 0541 1003 1004 1005
Böltken, F. 1337
Bötticher, K. W. 0918
Boh,K. 0272 0625 0807
Bollmus, R. 8028
Bonhöffer, M. 0553
Boskma, P. 0804
Boynton, G. R. 1540
Brachtl, W. 1141 1142 1241 1242
Brandes, U. 1595
Braun, J. 1663
Brauneck, A.-E. 0254
Brinkmann, G. 0180 0338 0369 0982
Brody, R. 1408
Bronfenbrenner, U. 0153
Brothun, M. 1481 1818
Brusten, M. 0872
Buchhofer, B. 0766 0767
Bühringer, G. 1337
Bürklin, W. P. 1221 1681
Bujok-Hohenauer, E. 1457
Burchards, E. 5006
Burg, F. van der 0804
Burger, A. 1234
Burgess, E. W. 0252
Butler, D. 1407
Cameron, St. H. 0920
Campo, S. del 0803
Capitani, F. de 8018 8019 8020
Chatterjee, B. B. 0810
Clark, S. M. 1409 1410
Coleman, J. S. 1196
Conradt, D. P. 0243
Converse, Ph. 0275 0276 0625 1402
Cotgrove, St. 1387
Cronenberg, P. 0359
Cseh-Szombathy, L. 0270 0625
Czaia, U. 1635
Dabier, M. 1126
Dähne, E. 0341
Daheim, H.-J. 0187 0355 0360 0412 0644
Dahl, R. A. 0237
Dallenbach-Hellwig, G. 1124
Dalton, R. J. 1090 1091
Dangschat, J. S. 1183 1685
Daniel, K. 0317
Deeley, E. C. 1377 1378
Demitch, B. 0403
Dennis, J. 1408
Deth, J. W. van 1188 1189
Deutsch, K. W. 0085 1511
Devereux, C. 0153
Diederich, N. 1780
Diekershoff, K.-H. 0394 0549 0678 0679 0895
Diekershoff, S. 0549 0678 0679 0895
Dierkes, M. 0306
Dittrich, G. G. 0551
Does, K.J. 0820
Dogs, W. 1124
Donsbach, W. 5014 5015 5016
Drop, G. 8054
Droth, W. 1183
Duncan, O. 0822
Ebel, H. 0267
Ebeling, D. 8014
Eberwein, W.-D. 1511
Eckart, C. 0821

Register der Erhebungsinstitute und Primärforscherinstitute

Neben dem Primärforscher-Register, welches allein die Namen der jeweiligen Personen angibt, informiert dieses Register über die Institutionen, in denen der Forscher tätig ist bzw. war. Ebenso sind die Erhebungsinstitute verzeichnet.

Forschungsinstitut für Soziologie, Universität zu Köln
0002 0036 0055 0056 0057 0058 0059 0060 0061 0062 0084 0086 0129 0130 0131 0153 0173 0195 0232 0233 0351 0352 0353 0364 0365 0366 0367 0393 0600 0639 0640 0813 0838 0862 0959 0960 0961 0996 1032 1539 2000
Forschungsstelle der Hochschule für Wirtschaft und Politik, Hamburg
0160 0526 0766 0767 0816
Forschungsstelle für empirische Sozialökonomik, Köln
0016 0082 0083 0094 0095 0096 0134 0136 0140 0165 0180 0306 0307 0308 0338 0369 0982 1182 1305
Forschungsstelle für gesellschaftliche Entwicklungen, Universität Mannheim
1805
Forschungsstelle für Jugendfragen, Hannover
0528
Forschungsstelle Jugendliche und randständige Gruppen, Köln
0873
Forschungsstelle Sozialökonomik der Arbeit, Freie Universität Berlin
1346 1348 1723 1724 1725 1726
Forschungsstelle vergleichende Stadtforschung am Institut für Soziologie, Universität Hamburg
1625
Forschungszentrum für schweizerische Politik, Universität Bern
8025 8026
Foundation for Child Development, New York
1170
Freie Universität Berlin
1053 1059 1061 1188 1189 1190 1191 1533 1534 1535 1537
Friedrich-Ebert-Stiftung, Bonn-Bad Godesberg
1731 1732 1733
Friedrich-Meinecke-Institut, Freie Universität Berlin
8044
Gallup International Research Institutes
0936
Gallup Markedsanalyse, Hellerup
1197 1318 1319 1320 1321 1541 1542 1543 1544 1684 1712 1713 1714 1715 1750
Gallup Markedsanalyse, Kopenhagen
0628 0970 0986 0987 0988 0989 0990 0991 0992 0993 0994 0995 1036 1037 1038 1039 1205 1206 1207 1208 1209
Gallup of Uruguay
0401
Genossenschaftliches Seminar der Universität zu Köln
0162
Gesamthochschule Duisburg
1184
Gesellschaft für Grundlagenforschung, München
0651 0681 0815 1326
Gesellschaft für Mathematik und Datenverarbeitung (GMD), Bonn
1312
GETAS, Bremen
0084 0757 0765 0783 0784 0785 0786 0787 0788 0789 0790 0791 0823 0839 0840 0841 0842 0851 0852 0853 0854 0855 0856 0857 0861 0895 1000 1017 1090 1091 1098 1099 1123 1124 1125 1126 1127 1137 1139 1152 1153 1154 1155 1156 1160 1188 1189 1190 1191 1213 1214 1231 1233 1236 1322 1335 1336 1337 1340 1391 1392 1393 1394 1395 1396 1397 1398 1399 1415 1416 1417 1418 1419 1420 1431 1433 1441 1452 1486 1502 1503 1504 1505 1509 1570 1571 1582 1601 1617 1681 1723 1724 1725 1726 1736 1737 1738 1790 1795 1807 1809 1810
GfK, Nürnberg
1001 1149 1150 1182 1243 1316 1426 1427 1428 1429 1618 1619 1704 1782
GfM, Gesellschaft für Marktforschung, Hamburg
0483 0484 0627 0628
GFM-GETAS, Hamburg
1487 1583 1584 1585 1586 1587 1588 1589 1590 1618 1619 1641 1670 1680 1705 1707 1708 1728 1782 1783 1804
Group of Social Prognoses, Institute of Philosophy and Sociology, Polish Academy of Sciences, Warschau
0809
Gruner + Jahr, Hamburg
1065 1066 1067 1068 1123 1124
Gruppe "UNESCO Sozialforschung in Deutschland", Sozialforschungsstelle an der Universität Münster, Dortmund
0254
Hessisches Institut für Bildungsplanung und Schulentwicklung, Wiesbaden
1435
Historisches Institut, Universität Bern
8018 8019 8020 8048
Historisches Seminar, Universität Bonn
8029
Historisches Seminar, Universität Münster
8016
Historisches Seminar, Universität zu Köln
8041
Hochschule für Sozial- und Wirtschaftswissenschaften, Linz
1462

Geographisches Register

Im geographischen Register sind alle Erhebungsregionen (z.B. Länder oder Städte) der Untersuchungen aufgeführt.

Niederlande 0025 0076 0078 0121 0143 0191 0421 0501 0626 0627 0628 0759 0765 0804 0970 0986 0987 0988 0989 0990 0991 0992 0993 0994 0995 1036 1037 1038 1039 1104 1106 1108 1110 1112 1188 1189 1197 1205 1206 1207 1208 1209 1287 1318 1319 1320 1321 1354 1405 1431 1511 1541 1542 1543 1544 1680 1684 1712 1713 1714 1715 1750 1752 1753 2000

Niedersachsen 0528 0703 0746 0877 1126 1244 1245 1518 1519 1611 1612

Nigeria 0149

Nima (Ghana) 5006

Nordamerika 0936

Norddeutscher Bund 8006 8008 8011

Nordirland 2000

Nordrhein-Westfalen 0013 0053 0062 0105 0106 0110 0137 0151 0152 0185 0233 0247 0253 0338 0357 0466 0467 0473 0495 0542 0549 0561 0562 0600 0639 0640 0702 0746 0817 0818 0864 0873 0996 1004 1015 1016 1032 1126 1179 1440 1441 1443 1444 1467 1468 1611 1612 1624 1794 1812

Osterreich 0642 0758 0765 1292 1364 1431 1490 1620 1680 2000 8022

Östringen 0461 0462

Offenbach 0459 1258 1262 1264

Oldenburg 0452 0453

Olomouc (Tschechoslowakei) 0269 0625

Oslo 0205

Osnabrück 0555

Osterwick 0582

Osmiederbayern 8032

Panama 0400

Paraguay 0402

Paris 0306

Pelarco (Chile) 5009 5026

Philadelphia (USA) 0170

Polen 0809 1159

Portugal 1300 1542 1543 1544 1712 1713 1714 1715 1750 1752 1753

Puerto Rico 0399

Queidersbach 0162

Quernheim 8031

Quilmana (Peru) 0278 0625

Erhebungszeitraum-Register

Dieses Register gibt Auskunft über den Zeitpunkt, zu dem die Datenerhebungsphase der jeweiligen Studie begonnen hat. Neben der chronologischen Einordnung der Studien, kann anhand dieses Indexes ein Überblick über das zeitliche Spektrum der im Zentralarchiv erhältlichen Datensätze gewonnen werden.

1970 0487 0491 0492 0514 0526 0541 0552 0606 0609 0610 0611 0612 0613 0626 0639 0641 0643 0649 0651 0655 0678 0702 0703 0739 0740 0741 0790 0791 0801 0846 0962 1031 1032 1102 1124 1159 1169 1358 1405 1538 1740 1741 1742

1971 0515 0516 0517 0518 0542 0547 0548 0549 0568 0601 0602 0603 0604 0607 0614 0615 0616 0617 0618 0627 0629 0642 0652 0654 0656 0664 0668 0681 0683 0742 0743 0769 0819 0821 0843 0847 0848 0863 0864 0868 0872 0877 0878 1103 1104 1410

1972 0598 0599 0605 0608 0619 0620 0621 0622 0623 0630 0631 0632 0633 0635 0636 0637 0638 0645 0646 0653 0665 0669 0670 0672 0679 0682 0685 0744 0745 0746 0747 0748 0749 0750 0751 0768 0796 0816 0828 0839 0840 0841 0849 0850 0873 0881 0951 0953 0982 1022 1023 1024 1105 1106 1123 1168 1199 1202 1403 1408 1448 1501

1973 0628 0648 0657 0658 0659 0660 0661 0662 0663 0666 0671 0676 0752 0760 0765 0770 0812 0815 0822 0829 0835 0842 0851 0858 0860 0865 0871 0910 0911 0912 0919 0939 0940 0954 0959 1006 1007 1008 1009 1011 1012 1013 1014 1015 1016 1021 1025 1065 1066 1067 1070 1088 1107 1108 1133 1322 1431 1502 1684

1974 0677 0754 0755 0757 0758 0759 0761 0797 0799 0817 0818 0830 0838 0852 0859 0866 0867 0874 0882 0883 0884 0885 0903 0913 0914 0915 0916 0918 0922 0936 0937 0944 0945 0960 0963 0986 1109 1110 1126 1182 1211 1232 1236 1514 1515 1812

1975 0756 0762 0763 0764 0771 0772 0798 0820 0831 0853 0869 0870 0886 0887 0888 0889 0895 0905 0917 0928 0930 0931 0932 0933 0948 0952 0955 0957 0976 0987 0988 0998 1026 1064 1082 1092 1111 1112 1127 1134 1271 1326 1458

1976 0800 0823 0832 0854 0861 0862 0906 0907 0908 0909 0923 0929 0967 0969 0980 0983 0989 0990 1001 1027 1034 1035 1042 1068 1071 1078 1079 1090 1093 1094 1095 1096 1097 1098 1113 1125 1135 1144 1145 1170 1184 1210 1233 1238 1239 1391 1392 1393 1394 1395 1396 1397 1457 1503 1605

1977 0814 0833 0855 0904 0924 0925 0949 0958 0968 0970 0971 0972 0991 0992 1002 1005 1019 1028 1030 1043 1044 1075 1085 1099 1114 1237 1369 1481 1504 1505 1559 1562

1978 0834 0856 0857 0965 0977 0978 0984 0993 0994 0995 1003 1004 1029 1040 1074 1076 1083 1084 1086 1089 1115 1137 1141 1143 1172 1175 1178 1179 1185 1220 1241 1254 1323 1362 1363 1415 1416 1417 1418 1420 1443 1444 1506 1560

1979 1017 1036 1037 1041 1045 1047 1052 1069 1072 1073 1077 1087 1091 1116 1128 1129 1146 1147 1173 1174 1176 1177 1181 1183 1188 1189 1224 1243 1259 1268 1269 1272 1313 1324 1341 1343 1354 1360 1361 1419 1447 1507 1508 1593 1701 1702 1703

1980 1000 1038 1039 1051 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1136 1142 1152 1164 1165 1166 1167 1187 1190 1191 1200 1203 1205 1221 1223 1225 1240 1242 1253 1258 1260 1261 1262 1263 1274

Stichwortverzeichnis

Den Hauptteil des Registers bildet ein ausführliches Stichwortverzeichnis.

Massenmedien 0314 0345 0372 0374 0443 0513 0641 1198 1199 1533
Masturbation 0325 1710
Matthöfer 1059 1063
Mauer 0066 0245 1666
Mauerbau 0066 0067 0245 1526
MBFR-Verhandlungen 0923 1511
Media-Analyse 0848 0849 0850 08510852 0853 0854 0855 0856 0857 1152 1153 1154 1155 1156 1617 1618 1619 1704 1705 1782 1783
Medien 0055 0133 0142 0208 0443 0454 0557 0560 0635 0897 0903 0908 0909 0921 0988 0991 1036 1040 1128 1139 1160 1170 1198 1200 1202 1204 1208 1214 1238 1312 1318 1319 1325 1340 1357 1377 1390 1400 1460 1526 1582 1587 1590 1597 1609 1635 1666 1670 1728 1730 1732 1733 1763 1809
Medienberichterstattung 0627 1198 1274 1443 1586 1587 1713 1760 1806
Mediennutzung 00010003 0012 0013 0024 0033 0037 0058 0059 0062 0064 01110113 0127 0128 0142 0143 0145 0147 0156 0163 0166 0179 0195 0231 0232 0236 0237 0238 0241 0246 0253 0266 0267 0268 0288 0289 0293 0296 0302 0303 0314 0326 0363 0375 0396 0406 0464 0465 0466 0478 0499 0516 0518 0519 0522 0523 0525 0528 0530 0532 0534 0536 0542 0549 0556 0557 0560 0564 0574 0581 0606 0608 0642 0654 0665 0702 0705 0738 0747 0788 0789 0844 0845 0846 0847 0848 0849 0850 0851 0852 0853 0854 0855 0856 0857 0867 0925 1001 1003 1005 1021 1038 1052 1068 1070 1087 1092 1098 1099 1128 1136 1152 1153 1154 1155 1156 1198 1200 1202 1204 1213 1214 1318 1390 1400 1404 1405 1406 1407 1408 1409 1434 1548 1570 1571 1582 1597 1617 1618 1619 1621 1637 1638 1639 1683 1704 1705 1729 1730 1750 1782 1783 1792 1809 1826
Medikament 0185 0334 0335 0336 0419 0578 0580 0770 0829 0868 0904 0906 0936 1001 1052 1068 1070 1087 1141 1210 1432 1433 1570 1627 1706 1826
Medikamentenkonsum 0868 0904 1124 1210 1339 1432 1433 1447 1548 1603 1710 1728
Meditation 0769
Medizin 0380 1166 1274 1367 1546 1728
Mediziner 1074 1728
Meer 0098 1166 1546 1719
Mehrparteiensystem 0066 0067 0328 0375 0382 1540
Meinungsäußerung 0521 0628 1061 1073 1224
Meinungsbildner 1632 1633 1634 1785
Meinungsbildung 0241 0628 0635 0792 1139 1198 1199 1325 1404 1405 1406 1407 1408 1589
Meinungsfreiheit 0375 0422 0530 0626 0991 1400 1490
Meinungsfühler 0011 0013 0061 0127 0128 0186 0208 0241 0243 0248 0280 0295 0297 0298 0299 0327 0380 0381 0382 0426 0438 0462 0466 0481 0484 0498 0517 0534 0562 0575 0627 0628 0747 0777 0778 0779 0782 0783 0784 0786 0839 0948 0989 0990 0991 0992 0993 0994 1017 1036 1037 1038 1039 1043 1070 1087 1090 1098 1099 1148 1149 1150 1161 1172 1174 1176 1201 1206 1207 1208 1209 1236 1318 1319 1320 1321 1323 1324 1365 1390 1426 1434 1449 1501 1506 1541 1542 1543 1544 1684 1712 1713 1714 1715 1722 1729 1750 1792 1809 1810 1811
Meinungsgefolgschaft 0012 0060 0438 0466 0575 0576 0747 0777 0778 0779 0782 0783 0784 0786 1070 1098 1099 1148 1150 1172 1174 1323 1324 1426 1506
Meinungsübereinstimmung 1173 1174 1175
Meinungsverschiedenheiten 0091 0523 0575 0577 0578 0579 0581 0746 0757 0772 0957 0959 0960 0968 1142 1339 1360 1559 1569 1613 1728
Meldepflicht 1642 1649 1662 1670 1698 1757 1758
Mende 0011 0043 0050 0314 0330 0556 0557
Mendes-France 0430 0431
Menschenrechte 0970 1090 1091 1159 1420 1715
Menstruation 1124 1210 1710
Messe 0483 0484
Mexiko 0030 0398 0624 09811521
Mietbeihilfe 0526 1169
Miete Olli 0134 0151 0160 0166 0170 0174 0198 0235 0253 0288 0297 0460 0489 0526 0551 0566 0606 0607 0655 0658 0666 0668 0671 0676 0684 0766 0816 0820 0821 0858 0860 0928 0939 0946 0947 0948 0952 1001 1002 1003 1005 1045 1065 1068 1098 1099 1137 1144 1157 1158 1169 1262 1270 1276 1305 1328 1358 1368 1455 1578 1579 1599 1604 1626 1629 1630 1631 1678 1736 1737 1738 1756 1786 1787 8002
Mieter 0332 0336 0337 0666 0671
Mietnebenkosten 1098 1099 16311786 1787
Mietpartei 1617 1618 1619 1685
Mietshaus 1377 1378
Mietverhältnis 0058 0059 0145 0155 0288 0550 0566 0581 0666 0668 0957 1635
Mietwagen 1166 1582 1722
Mietwohnung 0860 0863 1215 1722
Mietwucher 0581
Migration 0946 0947 1199 1411 1413 1414 8006 8007
Mikroelektronik 1274 1364 1367 1434 1582 1812
Mikrozensus-Erhebung 1583 1587 1590
Milch 0544 0577 0752 1373 1614 1637 1638 1639
Milieu 1580 1612 1729
Militär 0366 0404 04410457 1159 1715 2000
Militärdienst 0085 0160 0364 0457 0490
Militarismus 0178 1026 1027
Mineralölsteuer 1508
Minigolfanlage 1164
Ministerpräsident 0053 0442 0472 0562 0563 0702 0865 0871 1040 1079 1126 1172 1173 1174 1175 1223 1236 1244 1399 1466 1467 1503 1517 1518 1519 1555
Ministerpräsidentenpräferenz 1040 1223 1417 1419 1551 1651 1652 1654 1656 1657 1658 1661 1681 1697 1699
Ministerrat 1714
Mischehe 1513
Mißtrauensantrag 0245 0598 0665
Mitbestimmung 0107 0286 0312 0313 0370 0394 0395 0426 0514 0525 0530 0558 0628 0630 0631 0633 0751 0777 0788 0801 0819 0836 0859 0876

An das
Zentralarchiv für empirische Sozialforschung
an der Universität zu Köln
Bachemer Str. 40
D-5000Köln41

EARN / BITNET: BAUSKE @ DK0ZA1
Telefon: 0221/4 76 94-50
Telefax: 0221/4 76 94-44

Bestellung

des Datenbestandskatalogs 1990

Ich bestelle.....Exemplar(e):

Zentralarchiv für empirische Sozialforschung (Hg.)
Daten der empirischen Sozialforschung:
Datenbestandskatalog des Zentralarchivs mit
Beschreibungen von Daten der empirischen
Sozialforschung von 1945 bis 1990 und Daten
der historischen Sozialforschung.
Campus Verlag , Frankfurt/New York 1991

zum Einzelpreis von 48.- DM

Name: _____

Anschrift: _____

_____

Telefon: _____

Datum: _____

Unterschrift: _____

Das Zentralarchiv auf dem Soziologentag in Frankfurt

Vom 9. - 12. Oktober 1990 wurde in Frankfurt der 25. Deutsche Soziologentag veranstaltet. Nach 1910, 1946 und 1968 war dies der vierte Soziologentag, der in Frankfurt stattfand. Der Kongreß stand unter dem Thema "Die Modernisierung moderner Gesellschaften" und befaßte sich ganz generell mit Gegenwartsproblemen und Entwicklungstendenzen in modernen Gesellschaften. Obwohl schon lange vor den dramatischen Ereignissen um die Vereinigung der beiden deutschen Staaten festgelegt, paßte das Thema genau in die Zeit.

Neben den Vorträgen und den Veranstaltungen der Sektions- und Ad-hoc-Gruppen wurde Verlagen und Instituten die Möglichkeit zur Selbstdarstellung geboten, die auch vom Zentralarchiv wahrgenommen wurde. Selbst wenn die Präsentationsmöglichkeiten nicht so komfortabel und so zentral gelegen waren wie beim letzten Soziologentag in Zürich, so hat dennoch eine sehr große Zahl von Benutzern und Interessenten des Zentralarchivs unseren Stand aufgesucht.

Das Zentralarchiv stellte sein Angebot an Dienstleistungen für die Sozialforschung vor. Neben neuen Codebüchern - auch solchen der ersten DDR-Studien im Hause - wurde vor allem ein Vorabdruck des neuen Datenbestandskatalogs angeboten. In einer maschinenlesbaren Version dieses Katalogs waren auf Wunsch der Benutzer Recherchen an einem Personal Computer möglich. Hierfür wurde das Retrieval-System askSam eingesetzt. Zentralarchivbenutzer und Interessenten haben die Gelegenheit wahrgenommen, sich in zum Teil intensiven Gesprächen über neuere Datensätze zu informieren.

Anläßlich des Soziologentages in Frankfurt haben wir, wie bereits bei früheren Soziologentagen, eine Sonderpublikation erstellt. Neben einer kurzen Darstellung des Dienstleistungsangebotes des Zentralarchivs werden darin ausgewählte Datensätze vorgestellt, wie z.B. Studien im Rahmen des ISSP (International Social Survey Programme), der ALLBUS-Reihe sowie Datensätze aus der ehemaligen DDR und die Eurobarometer.

Sonderpublikation zum Soziologentag '90

Die Broschüre enthält außerdem eine kurze Beschreibung des Datenbankprojekts zu Entwicklungsländern, welches vom Zentralarchiv in Zusammenarbeit mit der Arbeitsstelle für Entwicklungsländerforschung der Universität zu Köln durchgefühlt wird. Darüber hinaus stellen wir in einem gesonderten Abschnitt die Dienstleistungen des ZHSF (Zentrum für Historische Sozialforschung) vor. Eine Übersicht zu den Veröffentlichungen und weiterem Informationsmaterial des Zentralarchivs schließt sich an.

Von den vorhandenen Restbeständen dieses 32-seitigen Heftes senden wir Ihnen auf Anfrage - die auch über BAUSKE at DK0ZA1 via EARN möglich ist - gerne ein Heft kostenlos zu.

ALLBUS: Eine Übersicht zu dem gegenwärtigen Datenangebot

I. Allgemeines

Der ALLBUS (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften) wurde Mitte der siebziger Jahre als gemeinsames Vorhaben des Zentrums für Umfragen, Methoden und Analysen e.V. (ZUMA, Mannheim), des Zentralarchivs für empirische Sozialforschung und einer Gruppe von Primärforschern konzipiert.[1]

Kernstück der ALLBUS-Konzeption sind regelmäßig zu wiederholende, repräsentative Bevölkerungsumfragen mit einem teils konstanten, teils variablen Fragenprogramm, das zentrale inhaltliche Bereiche der empirischen Sozialforschung abdeckt. Sie dienen der Erhebung und Verbreitung sozialwissenschaftlicher Daten, die in Forschung und Lehre allgemein verfügbar sein sollen. Seit 1980 wurden ALLBUS-Umfragen jeweils im Zweijahresabstand durchgeführt. Die einzelnen Datensätze umfassen unabhängig gezogene Stichproben aus der Bundesrepublik und West-Berlin mit einem Umfang von ca. 3000 deutschen Befragten, die mindestens 18 Jahre alt sind.

In jedem ALLBUS ist auch eine umfangreiche Standarddemographie enthalten, welche es u. a. ermöglicht, Sozialstrukturanalysen durchzuführen oder den sozialen Hintergrund der Befragten bei anderen Untersuchungen differenziert zu berücksichtigen.

Die ALLBUS-Erhebungen eröffnen auch solchen Interessenten, die nicht über eigene Primärdaten verfügen, den Zugang zu einem breitem Spektrum gesellschaftlich relevanter Fragestellungen. Mit den vorliegenden Querschnittsdaten können eine Vielzahl von Einstellungen und Verhaltensweisen der Bundesbürger detailliert analysiert werden. Die bereits vorliegenden ALLBUS-Zeitreihen ermöglichen darüber hinaus die Untersuchung sozialen Wandels.
Um die ALLBUS-Studien einer breiten Öffentlichkeit zugänglich zu machen, bietet das Zentralarchiv u. a. folgende Dienstleistungen an:

- kostengünstige Bereitstellung von ausführlichen Dokumentationen über die Datensätze
- Aufbereitung der Daten mit dem Ziel möglichst einfacher Zugänglichkeit für Auswertungen

1 Der Primärforschergruppe gehören zur Zeit folgende Personen an: Klaus *Allerbeck* (Universität Frankfurt), Karl Ulrich ***Mayer*** (Max-Planck-Institut für Bildungsforschung, Berlin), Walter *Müller* (Universität Mannheim), Karl Dieter ***Opp*** (Universität Hamburg), Franz Urban *Pappi* (Universität Mannheim, federführend), **Erwin K. *Scheuch*** (Universität zu Köln), Rolf *Ziegler* (Universität München). M. Rainer *Lepsius* (Universität Heidelberg) gehörte von 1980-1982 ebenfalls zu den Primärforschern.

- Distribution der Daten und Dokumentationen an die interessierte Öffentlichkeit
- Beratung bei Fragen zum Dateimanagement mit SPSS

Die ALLBUS-Studien sind gegenwärtig die in Deutschland am häufigsten verwendeten sozialwissenschaftlichen Datensätze.[2] Der ALLBUS gewährleistet somit eine kontinuierliche und inhaltlich fruchtbare Basis für Analysen, deren Ergebnisse von verschiedenen Anwendern wechselseitig leicht zu überprüfen oder weiterzuentwickeln sind.

Nachfolgend wird ein kurzer Gesamtüberblick über die bisher im Zentralarchiv verfügbaren ALLBUS-Einzelstudien (Abschnitt 2) und ALLBUS-Zusatzstudien (Abschnitt III) gegeben. Der Abschnitt IV stellt die beiden verschiedenen Versionen des kumulierten ALLBUS gegenüber, wobei besonderes Schwergewicht auf die Präsentation des neuen ALLBUS 1980-88 gelegt wird. Zusätzliche Informationen über die Datenweitergabe, Codebuchdokumentationen, Benutzungsgebühren und Zugangskategorien sind im Anhang I enthalten. Eine kurze Hilfestellung, wie die kumulierten ALLBUS-Daten auf dem PC eingesetzt werden können, ist im Anhang II enthalten.

Ansprechpartner im Zentralarchiv für Fragen zum ALLBUS sind:
Rolf *Uher* (Tel.: 0221 / 47694 - 25) und Michael *Terwey* (Tel.: 0221 / 47694 - 22).

2. ALLBUS-Einzelstudien

ALLBUS 1980 (ZA-Studien-Nr. 1000):
Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM
Das Codebuch enthält zusätzlich nationale und internationale Vergleichsdatenauszählungen.

Erhebungsschwerpunkte: Wichtigkeit von Lebensbereichen; Einstellungen und Kontakte zu Gastarbeitern; Wichtigkeit von Berufsaspekten; Erziehungsziele; Freundschaftsbeziehungen; Einschätzung sozialer Konflikte in der Bundesrepublik; Behördenkontakte sowie Einstellungen zu Behörden, zur Familie und zum Wohlfahrtsstaat.

---

2 Eine Bibliographie der bisher mit ALLBUS-Daten publizierten wissenschaftlichen Arbeiten ist bei ZUMA e.V. (Postfach 122155,6800 Mannheim 1) kostenlos erhältlich. Sie enthält zur Zeit 216 Publikationsnachweise mit kurzen Zusammenfassungen der Publikationsinhalte (Stand 30.04.90). Methodische Schwerpunkte stehen in 80 dieser Publikationen im Vordergrund, während sich 136 Arbeiten auf inhaltliche Themenstellungen konzentrieren.

ALLBUS 1982 (ZA-Studien-Nr. 1160):
Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 40 DM, Daten 100 DM
Das Codebuch enthält zusätzlich erweiterte nationale und internationale Vergleichsdatenauszählungen.

Erhebungsschwerpunkte: Wichtigkeit von Lebensbereichen; Wichtigkeit von Berufsaspekten; Erziehungsziele; gewünschte Eigenschaften von Kindern; Rolle von Frauen in der Familie; Einstellungen zur Abtreibung; Einschätzung der Wirtschaftslage und sozialer Konflikte in der Bundesrepublik; religiöse Einstellungen und religiöse Partizipation.

ALLBUS 1984 (ZA-Studien-Nr. 1340):
Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Fragen zur beruflichen Karriere; Relevanz von Eigenschaften für gesellschaftlichen Aufstieg; Beurteilung von Wohlfahrtsstaat und sozialer Ungleichheit; Einstellungen und Kontakte zu Gastarbeitern; Vertrauen zu öffentlichen Einrichtungen; Wahrnehmung von Umweltbelastungen; Einschätzung der Wirtschaftslage.

ALLBUS 1986 (ZA-Studien-Nr. 1500):
Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Wichtigkeit von Lebensbereichen; Einstellungen zur Abtreibung; Einschätzung der Wirtschaftslage; kulturelle Fertigkeiten; Erziehungsziele; Qualifikationsmerkmale der Befragten und Zeitangaben zum Qualifikationserwerb; persönliche Probleme und soziale Unterstützungsnetzwerke.

ALLBUS 1988 (ZA-Studien-Nr. 1670):
Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Einstellungen zum politischen System der Bundesrepublik und Formen politischer Partizipation; Fragen zum Thema AIDS; subjektiv empfundene Umweltbelastung; mögliche Gründe, als Deutscher stolz sein zu können; Einstellungen und Kontakte zu Gastarbeitern; Einstellungen zur Familie und den sich wandelnden Geschlechtsrollen.

Zusätzlich gibt es einen Vergleichsdatenband (ZA-Studien-Nr. 1727), in dem solche Variablen des ALLBUS 1988 im Zeitvergleich ausgezählt werden, die bereits früher im ALLBUS-Fragenprogramm enthalten waren (Gebühr 15 DM).

Zum Jahresanfang 1991 wird voraussichtlich auch der **ALLBUS 1990** an die Benutzer weitergegeben werden können. Dieser ALLBUS enthält neben einigen interessanten neuen Themenstellungen auch Schwerpunktreplikationen zum ALLBUS 1980, welche einen 10-Jahresvergleich ermöglichen.

ALLBUS 1990 (ZA-Studien-Nr. 1800):
Zugangskategorie A

Erhebungsschwerpunkte: Abweichendes Verhalten und Sanktion; Wichtigkeit von Lebensbereichen; Einstellungen zu Staat, Regierungsmaßnahmen, Behörden und sozialen Konflikten; politische Aktivitäten; Fragen zu den Themen AIDS und Abtreibung; Einstellungen zu Immigranten und Gastarbeitern; Kontakte zu Gastarbeitern; Einschätzung der Wirtschaftslage in der Bundesrepublik; Freundschaftsbeziehungen.

DL ALLBUS-Zusatzstudien

Für spezielle methodische Fragestellungen liegen ferner drei ALLBUS-Zusatzstudien vor. Die Zusatzstudie 1980 konzentriert sich auf die Erfassung von Variablen zur Analyse möglicher Interviewereffekte. In der Zusatzstudie 1984 wurde ein Teil der Befragten mit einem Variablensubset aus dem ALLBUS 1984 mehrfach befragt (Paneluntersuchung). Die Zusatzstudie 1986 hatte zum Hauptziel, diejenigen Befragten, bei welchen in der Hauptbefragung 1986 kein Interview realisiert werden konnte, in einer telefonischen Nachbefragung zu interviewen.

ALLBUS 1980 mit Methodenstudie (ZA-Studien-Nr. 1000):
Zugangskategorie C; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Diese Studie enthält als Ergänzung der weiter oben aufgeführten ALLBUS-Einzelstudie 1980 Daten mit Interviewermerkmalen und Befragungen der Interviewer zu einigen der im Hauptteil der Studie erhobenen Themen.

"Test-Retest"-Zusatzstudie zum ALLBUS 1984 (ZA-Studien-Nr. 1601):
Zugangskategorie C; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Einschätzung der Wirtschaftslage; Einstellungen zur sozialen Ungleichheit, zum Sozialstaat und zur Familie; politische Einstellungen; Einstellungen und Kontakte zu Gastarbeitern; demographische Angaben.

"Nonresponse"-Zusatzstudie zum ALLBUS 1986 (ZA-Studien-Nr. 1669):
Zugangskategorie C; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 25 DM, Daten 100 DM

Erhebungsschwerpunkte: Einschätzung der Wirtschaftslage; Fragen zum politischen Verhalten; Einstellungen zur Umfrageforschung und zum Datenschutz; demographische Angaben.

IV. ALLBUS-Kumulationsstudien

Um Zeitvergleiche und die Untersuchung relativ kleiner Befragtengruppen zu erleichtern, werden Daten aus den ALLBUS-Einzelstudien vom Zentralarchiv sukzessive in Kumulationsdateien zusammengefaßt. Für die Benutzer stehen die Kumulationsdaten in zwei unterschiedlichen Versionen zur Verfügung. Die erste Version, der kumulierte ALLBUS 1980-86 (ZA-Studien-Nr. 1335), beinhaltet Daten aus den vier ersten ALLBUS-Erhebungen (Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 30 DM, Daten 200 DM).

Gegenüber der früheren Kumulationsversion der Studie 1335 (1980-84), welche noch grundsätzlich alle Variablen aus den Einzelstudien umfaßte, wurde hier erstmals eine Reduktion vorgenommen. Die im Zeitverlauf wachsende Datenmenge machte für die neueren Versionen des kumulierten ALLBUS eine Entscheidung darüber notwendig, welche Daten künftig in der Kumulationsdatei weitergegeben werden sollen, weil ansonsten seitens der Benutzer Kapazitätsprobleme beim Einlesen und Auswerten der umfangreichen Kumulationsstudien auftreten können.[3]

Die Daten aus den Erhebungsjahren 1980-84 wurden in die Kumulationsdatei 1980-86 insgesamt übernommen, um so die bei der Kumulation erfolgte Detailüberarbeitung dieser Erhebungsjahrgänge an die ALLBUS-Benutzer weitergeben zu können. Diejenigen Variablen aber, die im ALLBUS 1986 erstmals erhoben wurden und vorherige ALLBUS-Variablen nicht unmittelbar ergänzen, bleiben von der Kumulation 1980-86 ausgeklammert. Hinter dieser neuen Konzeption steht der Grundgedanke, daß es nicht sinnvoll ist, die Kumulationsdatei künftig durch nur für eine ALLBUS-Erhebung geltende Informationen weiter zu vergrößern. Konkret würden andernfalls ca. 300 Variablen, die im Rahmen des ALLBUS-Gesamtprogramms 1986 das erste Mal erhoben wurden, die Kumulationsdatei belasten, ohne daß für sie vergleichbare Informationen aus früheren Erhebungsjahrgängen vorliegen. Dieses Problem würde mit jedem künftigen ALLBUS-Jahrgang, der in die Kumulation aufgenommen wird, weiter wachsen. Für solche 1986 erstmals enthaltenen Variablen können Analysen mit der Einzelstudie des ALLBUS 1986 schneller und einfacher gerechnet werden. Sollen für bestimmte Forschungsfragen dennoch Daten, die nicht im kumulierten ALLBUS 1980-86 enthalten sind, in eine spezifische Analysedatei aufgenommen werden, so kann dies mit dem SPSSx relativ einfach realisiert werden (vgl. das kommentierte Programmbeispiel in *Terwey* 1989: 34).

---

3 So referiert etwa *Wittenberg* 1988:117 f., daß Studenten in der Methodenlehre Schwierigkeiten hätten, ALLBUS-Einzelstudien, geschweige denn Kumulationsdateien, auf ihren Großrechner-Plattenbereichen zu halten. Nur Studenten, die das "handling" von SPSSx-Dateien gut beherrschten, könnten mit der Erstellung auswertungsspezifischer "system files" diese Kapazitätsprobleme lösen.

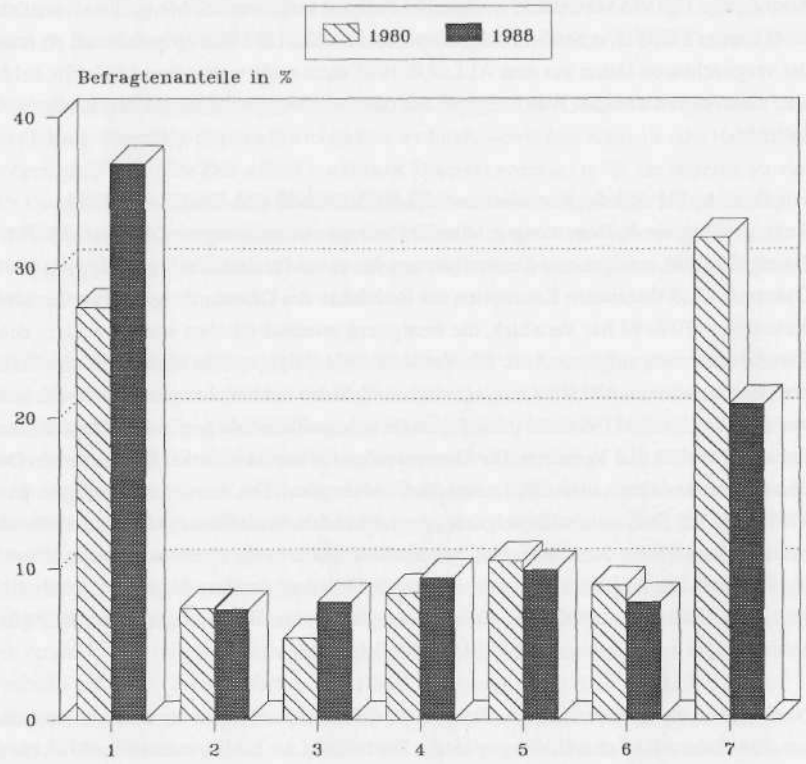
Der kumulierte ALLBUS 1980-86 umfaßt mit 658 Variablen im wesentlichen also **alle** Fragen aus den Einzelstudien 1980-84 und diejenigen Fragen aus dem ALLBUS 1986, welche bereits Replikationen aus dem vorherigen ALLBUS-Fragenprogramm sind. Die Daten wurden für die Kumulation in zahlreichen Details überarbeitet und durch neue, abgeleitete Variablen ergänzt. Benutzer, die mit den ALLBUS-Einzelstudien 1980, 1982 oder 1984 Analysen rechnen möchten, sollten deshalb in Erwägung ziehen, ob sie für diesen Zweck die Kumulation 1980-86 verwenden wollen. Die Fallzahl in dieser 13,3 Mb großen Kumulationsdatei beträgt 12045. Die Studie 1335 "Kumulierter ALLBUS 1980-86" wird nach Aufnahme der vergleichbaren Daten aus dem ALLBUS 1986 nicht mehr weiter entwickelt. Sie behält aber einen eigenständigen Wert aufgrund der darin enthaltenen Daten aus den Jahren 1980-84.

Für die **neue Version des kumulierten ALLBUS 1980-88** (ZA-Studien-Nr. 1795: Zugangskategorie A; Benutzungsgebühr für Hochschuleinrichtungen: Codebuch 30 DM, Daten 200 DM), erfolgte eine Generalisierung der zuvor für den Erhebungsjahrgang 1986 im Datensatz 1335 skizzierten Konzeption zur Reduktion des Datenumfangs: Es werden ab der Kumulation 1980-88 nur Variablen, die wenigstens zweimal erhoben worden sind, in den Benutzerdatensatz aufgenommen. Die Variablenreihenfolge, welche in den früheren Versionen des kumulierten ALLBUS aus technischen Gründen noch nicht optimiert werden konnte, wurde in der Studie 1795 unter inhaltlichen Gesichtspunkten neu gegliedert. Diese Kumulationsdatei enthält 466 Variablen. Die Daten wurden für die Kumulation in zahlreichen Details überarbeitet und durch neue, abgeleitete Variablen ergänzt. Die Anzahl der Befragten ist 15094. Bei der Datensatzgröße zeigt sich gegenüber dem kumulierten ALLBUS 1980-86 die positive Auswirkung der Konzeption, nur noch Variablen in die Kumulation aufzunehmen, die innerhalb des ALLBUS-Programms mehrfach erhoben wurden. Obgleich der kumulierte ALLBUS 1980-88 ca. 3000 Befragte mehr enthält, ist seine Datensatzgröße mit 12,3 Mb kleiner als im kumulierten ALLBUS 1980-86 (13,3 Mb).

Nachfolgend geben wir ein kurzes Beispiel für die Daten, welche in der Kumulation enthalten sind. Schaubild 1 enthält die graphische Darstellung der Häufigkeiten von V379 aus dem neuen kumulierten ALLBUS. Die Befragten wurden bei den Erhebungen 1980 und 1988 gebeten, anzugeben, ob sie dem Statement "Für terroristische Gewalttaten sollte die Todesstrafe eingeführt werden." zustimmend oder ablehnend gegenüberstehen. Insgesamt liegen 5781 valide Antworten vor.

In beiden Erhebungsjahren nehmen relativ viele Befragte extreme Haltungen zu dem Statement ein. Wurde aber 1980 nur in 27,4 % aller validen Antworten der Todesstrafe für terroristische Gewalttaten überhaupt nicht zugestimmt (Code 1), waren es 1988 bereits 36,9 %. Die Anteile der Befragten mit extremer Befürwortung des Statements (Code 7) fielen

Schaubild 1:
"Für terroristische Gewalttaten sollte die Todesstrafe eingeführt werden."



	1	2	3	4	5	6	7
1980	27,4	7,4	5,4	8,4	10,5	8,9	32
1988	36,9	7,3	7,8	9,4	9,9	7,8	20,9

Code 1: Stimme überhaupt nicht zu
Code 7: Stimme voll und ganz zu
Quelle: Kumulierter ALLBUS 1980-88

demgegenüber von 32 % (1980) auf 20,9 % (1988). Die zwischen den Extremen liegenden Fallzahlen zeigen nur geringfügige Veränderungen der relativen Größen. Insgesamt waren 1980 noch 51,4 % aller Befragten Befürworter der Todestrafe (Codes 5 bis 7), während 1988 mit 52 % aller Befragten sich eine Mehrheit gegen die Einführung der Todesstrafe aussprach (Codes 1 bis 3). Somit ergibt sich das Bild eines relativ starken Meinungsumschwungs, auf dessen Gründe wir hier nicht systematisch eingehen können. Ein Hintergrund des Meinungsumschwungs dürfte in der seit 1980 zurückgegangenen Gewalttätigkeit von Terroristen in der Bundesrepublik liegen. Dennoch mag die auch 1988 immerhin noch vorhandene Verbreitung von Akzeptanz der Todesstrafe vielen Beobachtern beunruhigend erscheinen. Die Sympathien mit dem politischen Terrorismus sind in der Bundesrepublik relativ niedrig, die Befürwortung der Todesstrafe als Sanktion ist dagegen relativ hoch (siehe den internationalen Vergleich von fünf Ländern in *Hewitt* 1990). Der direkte Datenzugriff auf relativ große Samples erlaubt uns, der Frage weiter nachzugehen, inwieweit sich bestimmte soziale Gruppen hinsichtlich ihrer Einstellung zur Todesstrafe unterscheiden.

In Tabelle 1 ziehen wir einige Merkmale aus der ALLBUS-Standarddemographie heran. Es zeigt sich, daß Frauen in der Bundesrepublik etwas seltener als Männer die Todesstrafe für terroristische Gewalttaten ablehnen (Frauen 44,7 %, Männer 47,5 %, Gesamtanteil der Ablehnenden 46 %).[4] Deutlicher ist die Differenzierung nach Bildungsabschlüssen und Alter. Personen, die maximal einen Hauptschulabschluß besitzen, sind mit 54,2 % häufiger für die strenge Sanktion als höher Qualifizierte (Gesamtanteil der Befürworter 45,1 %). Unter den jüngeren Befragten ist die Befürwortung der Todesstrafe signifikant seltener als unter den älteren (Altersgruppe 18-29 Jahre: 33,9 %). Hinsichtlich der Wahlabsicht für den Bundestag ergibt sich, daß CDU/CSU-Wähler relativ häufig für die Einführung der Todesstrafe sind (53,3 %). Sie werden allerdings darin von den NPD-Wählern noch übertroffen (78,3 %), soweit man für diese relativ kleine Gruppe der ALLBUS-Befragten (n = 23) eine Aussage treffen kann. Der Gegenpol wird von den potentiellen Wählern der Grünen gebildet, unter denen 71,7 % die Todesstrafe für terroristische Gewalttaten ablehnen. Die Gruppe der DKP-Wähler ist ähnlich stark in ihrer Ablehnung (85,7 %), doch ist hier die Fallzahl (n = 7) noch kleiner als bei der NPD-Wahlabsicht. Bemerkenswert ist schließlich, wie sich unterschiedlich religiöse Bundesbürger zu der Frage nach der Todesstrafe stellen. Gerade das Neue Testament hat ja die grundsätzliche Fragwürdigkeit der Todesstrafe an einem zentralen Fall aufgeworfen, der nach der Bibel nicht frei von politischen Konnotationen war. Die ALLBUS-Daten zeigen, daß die Gegnerschaft zur Todesstrafe signifikant nur unter den Konfessionslosen überrepräsentiert ist (61,1 %). Soweit die wenigen Angehörigen nicht-christlicher Religionsgemeinschaften uns eine Aussage erlauben (n = 14), ist auch hier eine Gegnerschaft

4 Diese Ausführungen beziehen sich auf Ausweitungen, die der Kürze halber nicht zwischen den beiden Erhebungsjahren differenzieren. Allgemein sei angemerkt, daß die dargestellten Unterschiede 1980 zumeist deutlicher ausfallen, daß aber auch 1988 die skizzierten Vergleiche in der Tendenz zutreffen.

Tabelle 1: Einstellung zur Einführung der Todesstrafe für terroristische Gewalttaten
(Berechnungen nach dem kumulierten ALLBUS 1980-88, Befragte aus den Erhebungsjahrgängen 1980 und 1988)

		Gegner %	Unentschiedene %	Befürworter %	n
Geschlecht	Frauen	44,7	9,5	45,8	3154
	Männer	47,5	8,1	44,3	2627
Ausbildung:	Maximal Hauptschulabschluß	36,2 *	9,6	54,2 *	3415
	Mittlere Reife	51,6 *	8,4	40,1 *	1255
	Hochschulreife	69,1*	6,9	24,0*	534
	Hochschulabschluß	70,8*	7,6	21,6*	510
Alter:	18 - 29 Jahre	58,6*	7,5	33,9*	1341
	30-44 Jahre	48,4	10,0	41,6	1554
	45 - 59 Jahre	41,1 *	9,1	49,7 *	1349
	60-74 Jahre	37,8*	9,3	52,9*	1198
	Über 74 Jahre	32,8*	7,8	59,4*	335
Wahlabsicht:	CDU/CSU	37,1 *	9,6	53,3 *	1627
	SPD	48,8	9,4	41,8 *	1874
	FDP	52,3	5,8	41,9	377
	NPD	13,0 *	8,7	78,3 *	23
	DKP	85,7	14,3	0,0	7
	Grüne	71,7*	5,5	22,8*	311
	Andere Partei	45,0	0,0	55,0	20
	Würde nicht wählen	42,5	7,5	50,0	254
Konfession:	Evangelische ohne Freikirchen	46.1	9,3	44,6	2588
	Evangelische Freikirchen	36,2	6,0	57,7 *	149
	Katholiken	43,7	9,1	47,2	2526
	Andere christliche Konfession	46,5	7,0	46,5	43
	Nicht-christliche Religionen	78,6	0,0	21,4	14
	Konfessionslose	61,1*	6,6	32,2*	422
Kirchgang:	Mehr als einmal in der Woche	43,8	8,6	47,7	128
	Einmal in der Woche	38,4 *	8,8	52,8 *	748
	Ein- bis dreimal im Monat	41,9	9,7	48,4	626
	Mehrmals im Jahr	42,3	10,9	46,8	1401
	Seltener oder nie	49,1 *	8,0	42,9 *	2382
Alle Befragten:		46,0	8,9	45,1	5781

* = Signifikanz der Abweichung von den statistisch erwartbaren Häufigkeiten < .01

Die Befragten mit den Codes 1 bis 3 aus Schaubild 1 wurden als Gegner der Todesstrafe zusammengefaßt, die Befragten mit den Codes 5 bis 7 als Befürworter. Prozentwerte in der Tabelle summieren sich zeilenweise zu 100 %. Die Kirchgangshäufigkeit wurde nur bei Mitgliedern christlicher Konfessionen erfragt.

weit verbreitet (78,6 %). Unter den Befragten, die Mitglieder einer traditionellen christlichen Konfession sind, weichen die Angehörigen der evangelischen Freikirchen am deutlichsten von der Gesamtheit ab - sie sind relativ häufig für die strenge Sanktion (57,7 %). Evangelische ohne Freikirchen (44,6 %) und Katholiken (47,2 %) liegen annähernd beim Gesamtprozentsatz der Todesstrafenbefürworter. Ziehen wir darüber hinaus die im ALLBUS nur bei christlichen Konfessionsmitgliedern erfragte Kirchgangshäufigkeit heran, so zeigt sich, daß gerade unter den sehr seltenen Kirchgängern und christlichen Konfessionsmitgliedern ohne Kirchgang die Gegnerschaft zur Einführung der Todesstrafe mit 49,1 % signifikant überrepräsentiert ist.

Die Diskussion demographischer Hintergründe zu der in Schaubild 1 dargestellten Einstellungsvariable sollte exemplarisch das für Analysen zur Verfügung stehende Datenpotential verdeutlichen. Sie mußte an dieser Stelle aus Platzgründen notwendigerweise kursorisch bleiben, so daß Zeitvergleiche, einige interessante Details und eine multivariate Analyse ausgeklammert wurden.

Anhang I:

Die in den Abschnitten II-3 genannten **Benutzungsgebühren** erhöhen sich für Anwendungen außerhalb des Hochschulbereichs bei den Einzel- und Zusatzstudien auf 200 DM für die Daten. Für die Kumulationsdaten (s. Abschnitt IV) erhöhen sich die Gebühren auf 300 DM. In begründeten Ausnahmefällen kann für Studenten eine reduzierte Leihgebühr vereinbart werden.

Alle aufgelisteten **Datensätze** werden standardmäßig auf Magnetbändern mit einem ausführlichen SPSS9- und einem SPSSx-Setup angeboten. Die Magnetbänder sind nach dem Überspielen der Daten an das Zentralarchiv zurückzugeben. Alternativ können die Magnetbänder auch in Rechnung gestellt werden. Die Datensatzgröße der ALLBUS-Einzelstudien liegt bei 2 bis 3 Mb.

Für SPSS-PC sind die genannten ALLBUS-Studien auch als SPSS-Exportfiles auf Disketten (Format 3 1/2" (1,44 Mb) oder 5 1/4" (1,2 Mb) lieferbar. ALLBUS-Kumulationsdateien werden für den PC-Einsatz in mehrere Teile zerlegt und können vor Ort mit SPSS-PC zusammengeführt werden (vgl. Anhang II). Die für die Datenübertragung erforderlichen Disketten müssen zusätzlich in Rechnung gestellt werden (pro Diskette 3 DM). Werden die Daten zugleich auf Band und Diskette bestellt, fällt zusätzlich eine Bearbeitungsgebühr von 10 DM an. Die Datensatzgröße der Exportfiles für Einzelstudien beträgt ca. 3 bis 4 Mb.

Die **Datendokumentationen** liegen in ZA-Codebüchern vor, welche u. a. Volltexte der Fragestellungen und Antwortkategorien, Randauszählungen sowie erklärende Anmerkungen enthalten. Den Codebüchern der ALLBUS-Einzelstudien ist außerdem eine Kopie der von ZUMA über die Erhebung erstellten Methodenberichte und eine Kopie des Fragebogens beigefügt.

Die Studien in der **Zugangskategorie** A sind allgemein für wissenschaftliche Arbeiten verfügbar. Nur für die Datensätze in der **Zugangskategorie** C ist es notwendig, daß vom Zentralarchiv eine schriftliche Genehmigung des Datengebers unter Angabe des Benutzers und des Auswertungszwecks eingeholt wird. Die **Codebücher** der ALLBUS-Studien in der Zugangskategorie C sind dagegen für wissenschaftliche Arbeiten frei verfügbar.

Anhang II:

Nachfolgend soll ALLBUS-Benutzern, die mit dem SPSS-PC bisher weniger vertraut sind, eine kurze Hilfestellung gegeben werden. Als Beispiel verwenden wir den neuen kumulierten ALLBUS 1980-88 (Studie 1795). Diese Daten liegen für die PC-Anwendung in Form von Jahrgangsspezifischen Exportfiles vor. Bevor Analysen gerechnet werden können, müssen die Exportfiles einzeln wie folgt importiert werden:

```
import file="sl795a.exp"
 /keep v2 v4 v5 v7 v8 v247 v466.
save outfile="sl795a.sys"
 /comp,
fin.
```

Der in diesem Beispielprogramm zum Import aufgerufene SPSS-PC-Exportfile heißt "sl795a.exp". Er enthält die Daten des ersten ALLBUS-Erhebungsjahres 1980 aus der Kumulationsdatei (Größe: ca. 3,2 Mb). Der Exportfile sollte nach unseren Erfahrungen in dem gleichen PC-Verzeichnis stehen wie das Importprogramm. Die folgende /keep-Anweisung selektiert für den Import sieben Variablen. Anwender mit der Programmversion 2 des SPSS-PC können hier bis zu 200 Variablen auflisten; Anwender mit der Programmversion 3 können, falls gewünscht, bis zu 500 Variablen auflisten - d.h. hier: Alle 466 Variablen des Kumulationsjahrgangs 1980 sind ggfs. in einem Lauf importierbar. Analoge Importprogramme sind beim Einlesen der Kumulation für alle bei den späteren Analysen involvierten Jahrgänge durchzuführen. Hierfür müssen im ersten Programmbeispiel nur die Dateinamen abgeändert werden (z.B. anstelle von "S1795a.exp" (Erhebungsjahrgang 1980) "sl795b.exp" für den Erhebungsjahrgang 1982). Der im ersten Programmbeispiel dargestellte Import benötigt auf einem IBM PS70 ca. 18 Minuten, der importierte Systemfile mit 7 ALLBUS-Variablen ist ca. 0,13 Mb groß.[5]

---

5 Analog ist die Vorgehensweise bei den ALLBUS-Einzelstudien. Diese liegen in dem neu überarbeiteten PC-Datenangebot jeweils in Form eines einzigen Exportfiles vor. Die nachfolgend dargestellte Prozedur "join add" kann deshalb entfallen. Für die Kumulation ist die Zerlegung in mehrere Subsetfiles u.a. deshalb sinnvoll, da sonst die Benutzer vor Ort die gesamte Kumulationsdatei geschlossen auf ihre Festplatte bringen müßten, was in jedem Falle ca. 17 Mb freien Speicherplatz erfordern würde. Werden für eine Analyse nicht alle Erhebungsjahrgänge benötigt, wie bei unserem Datenbeispiel in Schaubild 1, brauchen nur die jeweils relevanten Export- oder Systemfiles aktiviert zu werden.
Bedauerlicherweise kann SPSS-PC nur einen "missing value" als solchen behandeln. Da die meisten ALLBUS-Variablen mehr als einen "missing value code" enthalten, werden für die ALLBUS-PC-Exportfiles alle Definitionen fehlender Werte aufgehoben, um die ursprüngliche Codevielfalt zu erhalten. Eine analysespezifische Neudefinition fehlender Angaben kann vor Ort neu durchgeführt werden.

Im folgenden Beispielprogramm werden die jahrgangsspezifischen Systemfiles mit der Prozedur "join add" zusammengefügt:

```
get file="sl795a.sys".
join add file=*
 /füe=Msl795b.sys".
join add file=*
 /file="sl795c.sys".
join add file=*
 /file="sl795d.sys".
join add file=*
 /file="sl795e.sys".
crosstabs v5 v7 v247 v466 by v2.
save outfile="sl795add.sys" / comp,
fin.
```

Aufgerufen wird zunächst der im ersten Programmbeispiel erstellte Systemfile mit den Daten für 1980. Daran wird in den Zeilen 2 und 3 der Systemfile für den Erhebungsjahrgang 1982 (sl795b.sys) angehängt. Das Verfahren wiederholt sich bis zu den Daten des Erhebungsjahrgangs 1988 (sl795e.sys). Alle fünf ALLBUS-Jahrgänge Hegen nach diesem Programmabschnitt in einer vorläufig noch temporären Datei für Auswertungen vor. Die Prozedur "crosstabs" erstellt eine Kontrollauszählung für vier Variablen, aufgegliedert nach den in v2 abgespeicherten ALLBUS-Erhebungsjahrgängen. In der vorletzten Programmzeile wird, falls gewünscht, ein neuer komprimierter Systemfile (sl795add.sys) mit den Daten aus fünf Jahrgängen herausgeschrieben, der in unserem Beipiel ca. 0,62 Mb beansprucht. Die Gesamtlaufzeit des zweiten Beispielprogramms auf einem IBM PS70 beträgt ca. 12 Minuten. Weitere Hinweise zum Dateimanagement mit SPSS-PC können u.a. *Uehlinger* et al. 1988 entnommen werden. Dort werden in den Programmbeispielen ebenfalls Anwendungen mit ALLBUS-Dateien gezeigt.

Literatur:

Hewitt, Christopher 1990:
Terrorism and Public Opinion: A Five Country Comparison, in: Terrorism and Political Violence Vol. 2/2: 145-170.

Terwey, Michael 1989:
Der kumulierte ALLBUS 1980-86, in: ZA-Information 25: 30-35.

Uehlinger, Hans-Martin et al. 1988:
SPSS/PC+ Benutzerhandbuch (Bd. 1), Stuttgart, New York: Gustav Fischer.

Wittenberg, Reinhard 1988:
Methodenlehre mit ALLBUS-Daten: Erfahrungen aus Nürnberg, in: Soziologie. Mitteilungsblatt der Deutschen Gesellschaft für Soziologie 2/88: 99-128.

Michael Terwey

Das Thema AIDS in den Datensätzen des Zentralarchivs. Eine Bestandsaufnahme.

1. Einleitung

Seit dem Bekanntwerden von AIDS wurden weltweit wissenschaftliche Anstrengungen unternommen, um die Ursachen dieser Krankheit zu bekämpfen. Obwohl sehr schnell erkennbar war, daß mit AIDS weitreichende Folgen für das soziale Zusammenleben der Menschen in der Gesellschaft verbunden sind, wie z.B. die Veränderung des Sexualverhaltens weiter Bevölkerungsschichten sowie die Stigmatisierung und Diskriminierung von sozialen Minderheiten, wurde das Thema von der sozialwissenschaftlichen Forschung erst zögerlich aufgegriffen. Insbesondere empirisch angelegte Untersuchungen, die darauf abzielen, eine breite Datenbasis zur Beobachtung der gesellschaftlichen und sozialen Folgen von AIDS bereitzustellen, wurden in Deutschland vor 1987 recht selten unternommen.

Inzwischen hat sich die Situation jedoch gewandelt: Das AIDS-Zentrum im Bundesgesundheitsamt bemüht sich aktiv um die Einbeziehung der Sozialwissenschaften in die AIDS-Forschung. Hierzu werden eine Reihe von Dienstleistungen zur Verfügung gestellt, wie beispielsweise die Organisation von Arbeitstreffen sowie die Erstellung und Pflege einer AIDS-Projektdatenbank.[1] Auch das BMFT unterstützt mit seiner Forschungsförderung immer mehr Projekte der sozialwissenschaftlichen AIDS-Forschung. 1990 beträgt dessen Fördervolumen für sozialwissenschaftliche Projekte zum Thema AIDS fast 10% des Gesamtfördervolumens für die AIDS-Forschung (ca. 25 Millionen DM).[2]

Auch das Zentralarchiv kann als Serviceinstitut die empirische AIDS-Forschung unterstützen. Die Sammlung, Aufbereitung und Bereitstellung von Umfragedaten leistet einen Beitrag zur Verbesserung der Datenlage, insbesondere wenn mit Hilfe von Sekundäranalysen und Trenduntersuchungen die gesellschaftlichen Implikationen von AIDS im langfristigen zeitlichen Wandel beschrieben werden sollen.

---

1 Vgl. dazu z.B. *Herrmann., C; Heckmann, W.; Hessling, A.* (Hrsg.): AIDS und die Sozialwissenschaften: Dokumentation des 1. Arbeitstreffens im AIDS-Zentrum. Berlin: AIDS-Zentrum im Bundesgesundheitsamt (Selbstverlag), 1990, (AZ-Hefte: 1/1990) sowie *Herrmann., C; Heckmann, W.; Hessling, A.* (Hrsg.): AIDS und die Sozialwissenschaften: Eine Projektdatenbank deutschsprachiger Forscher/innen. Berlin: AIDS-Zentrum im Bundesgesundheitsamt (Selbstverlag), 1990, (AZ-Hefte: 2/1990).

2 Vgl. Bundesminister für Forschung und Technologie (Hrsg.): AIDS: Eine Herausforderung an die Wissenschaft 4. Bericht des Bundesministers für Forschung und Technologie. Bonn, 1990, S. 184 ff.

Im Zentralarchiv sind drei Datensätze archiviert, die sich intensiv mit dem Thema AIDS und den gesellschaftlichen Folgen befassen. Diese recht geringe Anzahl spiegelt die eingangs beschriebene Situation der sozialwissenschaftlichen AIDS-Forschung in Deutschland wider. Es sind jedoch weitere acht Mehrthemen-Umfragen archiviert, die sich neben anderen Fragen bzw. Fragenkomplexen auch mit AIDS beschäftigen.[3] Diese Datensätze werden im folgenden entweder einzeln oder - wenn möglich - summarisch in Form von Kurzbeschreibungen erläutert.[4] Ausgewählte Ergebnisse werden in Tabellenform dargestellt. Sie sollen einige Verwendungsmöglichkeiten der Daten in studienübergreifender Perspektive deutlich machen.

2. Datensätze zum Thema AIDS

2.1 Bundesweit repräsentative Datensätze der Bundeszentrale für gesundheitliche Aufklärung, Köln

Die Bundeszentrale für gesundheitliche Aufklärung (BZgA) begann frühzeitig mit der Beobachtung der gesellschaftlichen Folgen von AIDS anhand repräsentativer Bevölkerungsumfragen. Die Ergebnisse fließen direkt in die von der BZgA mitkonzipierten Informations- und Aufklärungskampagnen ein. 1987 und 1988 wurden im Auftrag der BZgA von FORSA zwei bundesweite Telefonumfragen erhoben, die zu einem großen Teil identische bzw. vergleichbare Frageformulierungen enthalten, so daß damit analysiert werden kann, inwieweit sich im Verlaufe eines Jahres Änderungen im Informationsstand, in der Bereitschaft zum Schutz vor AIDS und in den Einstellungen zu HlV-Infizierten ergeben haben.

Auch im Rahmen der Dauerbeobachtung des allgemeinen Gesundheitsverhaltens der Bevölkerung setzt die BZgA Umfragen ein. Die sogenannten "Effizienzkontrollen" sind im mehrjährigen Turnus durchgeführte Repräsentativbefragungen der deutschen Bevölkerung, die aus Modulen mit identischen Frageformulierungen bestehen. Die Ergebnisse dienen als Aktionsgrundlagen für die Arbeiten der BZgA. In die Effizienzkontrolle 1987 wurde erstmals ein Block mit Fragen zum Thema AIDS aufgenommen.

---

[3] Darüber hinaus sind in den Datenbeständen zwei weitere Studien zum Sexualverhalten enthalten, die in einigen Fragen den Einfluß von AIDS auf das Sexualleben thematisieren. Diese Untersuchungen sind unter den Archiv-Nummern 1798 (Das Leben zu zweit - Sexualverhalten 1988) sowie 1818 (Einstellungen zu Sexualität und AIDS, 1988) archiviert

[4] Ausführliche Studienbeschreibungen mit detaillierteren inhaltlichen Angaben, z.B. zu den Fragen, die über das Thema AIDS hinaus gehen, werden auf Anfrage kostenlos zugesandt.

Datensatzbeschreibung

AIDS im öffentlichen Bewußtsein der Bundesrepublik (April 1987 und Juli/August 1988)[5], Archiv-Nrn. 1758 und 1757

Datenerhebung: FORSA, Dortmund; 2013 bzw. 2006 Befragte

Grundgesamtheit und Auswahl

Untersuchungsgebiet: BRD einschl. West-Berlin. Die Auswahl der zu befragenden Personen erfolgte nach einem Zufallsverfahren, bei dem computergenerierte Zufallszahlen als Telefonnummern verwendet wurden, mit anschließender Zufallsauswahl der jeweils im Haushalt zu befragenden Personen. Befragt wurden Bundesbürger ab 16 Jahren (computergestützte telefonische Befragung).

Inhalt

Wissen über AIDS, Bereitschaft zum Schutz vor AIDS, Einstellung zu HIV-Infizierten. *Die Themen im einzelnen:* Zufriedenheit mit dem eigenen Gesundheitszustand; größte Gesundheitsgefahren für die deutsche Bevölkerung; gefährlichste Krankheiten in der Bundesrepublik; Kenntnis der Krankheit AIDS; Wissen über AIDS; Nutzung von Informationsquellen über AIDS; Einschätzung der von AIDS ausgehenden Gefahr; Einschätzung der Ansteckungsgefahr von AIDS in ausgewählten Situationen (Skala); vermutete Anzahl von HIV-Infizierten in der Bundesrepublik; Kenntnis der Abkürzung für den AIDS-Erreger; Durchführung eines HIV-Tests; gewünschte bzw. erhaltene Beratungsleistungen vor der Durchführung eines HIV-Tests; Kenntnis der Aussagekraft eines HIV-Tests; Kenntnis der Ansteckungsgefahren und vermutete Risikogruppen; eigener Kontakt zu AIDS-Kranken bzw. HIV-Infizierten; Verhaltensänderungen bzw. Veränderungen im täglichen Umgang mit Menschen aufgrund der AIDS-Gefahr; Sorge vor einer eigenen AIDS-Erkrankung bzw. einer AIDS-Erkrankung im Bekanntenkreis; Einschätzung der eigenen Infektionsgefahr aufgrund von Geschlechtsverkehr; Teilnahme an einer AIDS-Beratung sowie Beratungsort; Kenntnis der AIDS-Telefonberatung; subjektiver Informiertheitsgrad bezüglich AIDS; Gespräche im Freundes- und Bekanntenkreis über AIDS und mögliche Schutzmaßnahmen; Gespräche mit dem eigenen Partner über Schutzmaßnahmen beim Geschlechtsverkehr; Erwartung einer neuen Intimbeziehung in den nächsten Monaten und Absicht, mit dem neuen Partner über das Thema AIDS zu sprechen; empfundene Belastung einer Partnerschaftsbeziehung durch Gespräche über den Schutz vor AIDS bzw. durch die Durchführung eines HIV-Tests; sexuelle Aktivität; wechselnde Geschlechtspartner in den letzten 12 Monaten; Verhalten beim ersten Geschlechtsverkehr in einer neuen Liebesbeziehung; Alter

---

[5] Veröffentlichungen:
FORSA: AIDS im öffentlichen Bewußtsein der Bundesrepublik. Eine Wiederholungsbefragung im Auftrag der Bundeszentrale für gesundheitliche Aufklärung, Köln. Endbericht. Dortmund: Selbstverlag 1988.
Bundeszentrale für gesundheitliche Aufklärung: AIDS im öffentlichen Bewußtsein der Bundesrepublik. Wiederholungsbefragung. Zusammenfassung wichtiger Ergebnisse. Köln: Selbstverlag 1988.

beim ersten Geschlechtsverkehr; Einstellung zur Verwendung von Kondomen und zu anderen Schutzmaßnahmen; eigene Verwendung von Kondomen beim Geschlechtsverkehr sowie Häufigkeit der Kondomverwendung; Gründe der Kondomverwendung; Geschlechtsverkehr mit "unbekannten" Sexualpartnern und Kondomverwendung; Konsequenzen der AIDS-Gefahr für das eigene Sexualverhalten; Einstellung zu einer offenen Berichterstattung über mögliche Schutzmaßnahmen; Einstellung zur namentlichen Meldepflicht und vermutete Konsequenzen bei Einführung einer solchen Meldepflicht; Einstellung zur gesetzlichen Verpflichtung zu HIV-Tests; Bereitschaft zur finanziellen Unterstützung von AIDS-Beratungs- und AIDS-Hilfsinstitutionen; eigene Bereitschaft zur Betreuung von HIV-Infizierten; Einstellung zur Isolierung von AIDS-Kranken; vermuteter Einfluß der AIDS-Problematik auf die Vorurteile gegenüber Homosexuellen; Einfluß von AIDS auf die Sexualmoral; subjektiver Informiertheitsgrad über Schutzmaßnahmen; Wunsch nach zusätzlichen Informationen über das Thema AIDS; Zusammenleben mit einem Partner.

Ausgewählte Ergebnisse

Die Bedeutung von AIDS für die Gesellschaft zeigt sich deutlich darin, daß fast ein Viertel der Befragten sich schon einmal Sorgen gemacht hat, an AIDS zu erkranken. Dabei ist die Besorgnis besonders hoch bei denjenigen Befragten, die ohne Partner leben. Die meisten Sorgen machen sich Frauen bis 55 Jahre, die ohne Partner leben. Eine Gegenüberstellung der Ergebnisse der beiden BZgA-Umfragen verdeutlicht die Veränderungen innerhalb eines Jahres (vgl. Tabelle 1). Gesicherte Aussagen über die zeitliche Entwicklung der Besorgnis in der Bevölkerung, die in Relation zu gesellschaftlichen Parametern (z.B. Wirkungen von Aufklärungskampagnen, Stand der medizinischen AIDS-Forschung) gesetzt werden könnten, erfordern jedoch eine Fortsetzung der Befragungsreihe über einen längeren Zeitverlauf.

Eine weitere gesellschaftspolitisch bedeutsame Fragestellung ist, inwieweit die Bevölkerung die von staatlicher Seite diskutierten Maßnahmen, wie z.B. die namentliche Meldepflicht von HIV-Infizierten an das Gesundheitsamt sowie die Isolierung von AIDS-Kranken, bewertet. Die Einstellungen dazu sind in der Tabelle 2 für die beiden Befragungszeitpunkte wiedergegeben.

Tabelle 1: Sorgen vor einer AIDS-Erkrankung
(Angaben in Prozent)[6]

	1987	1988
Insgesamt	23	24
Männer	23	24
Frauen	23	24
Männer bis 55 Jahre	28	28
Frauen bis 55 Jahre	31	32
Männer bis 55 Jahre ohne Partner lebend	35	37
Frauen bis 55 Jahre ohne Partner lebend	39	41
Männer bis 55 Jahre mit Partner lebend	24	26
Frauen bis 55 Jahre mit Partner lebend	27	30

Tabelle 2: Einstellung zu staatlichen Maßnahmen im Rahmen der AIDS-Bekämpfung
(Angaben in Prozent)[7]

	1987	1988
namentliche Meldepflicht		
richtig	50	47
nicht richtig	39	42
weiß nicht	10	11
Isolierung von AIDS-Kranken		
richtig	15	6
nicht richtig	78	91
weiß nicht	6	3

[6] Frage: "Haben Sie sich auch schon einmal Sorgen gemacht, ob Sie selbst an AIDS erkranken könnten, oder bisher noch nicht?" Vgl. auch FORSA, 1988, a.a.O., S. 38 f.

[7] Fragen: "Halten Sie es für richtig oder nicht richtig, wenn alle diejenigen, die sich mit AIDS angesteckt haben, namentlich beim Gesundheitsamt gemeldet werden?" sowie "Halten Sie es für richtig oder nicht richtig, wenn man dafür sorgt, daß alle AIDS-Kranken mit niemandem sonst in Berührung kommen - mit Ausnahme des medizinischen Personals und der Angehörigen?" In jeweils weniger als 1% der Fälle wurde die Antwort verweigert.

Datensatzbeschreibung

Aktionsgrundlagen 1987 der Bundeszentrale für gesundheitliche Aufklärung; Effizienzkontrolle (Oktober 1987 bis November 1987), Archiv-Nr. 1728

Datenerhebung: GFM-GETAS, Hamburg; 1942 Befragte

Grundgesamtheit und Auswahl

Untersuchungsgebiet: BRD einschl. West-Berlin. Mehrstufig geschichtete Zufallsauswahl (ADM-Mastersample) von Personen im Alter von 14 Jahren und älter, die in Privathaushalten leben. Mündliche und schriftliche Befragung mit standardisiertem Fragebogen.

Inhalt: Informationsquellen zum Thema AIDS; Interesse an Informationen zu AIDS; Kenntnis und Wahrnehmung der AIDS-Aufklärungskampagne; Einschätzung der Vertrauenswürdigkeit ausgewählter Beratungsstellen und Medien bezüglich ihrer Informationen zu AIDS; eigene Vorschläge zur Bekämpfung von AIDS und zur Betreuung von AIDS-Kranken; Sorge, daß der Befragte sowie Freunde und Bekannte sich mit HIV infizieren könnten; eigene Verhaltensänderungen aus Furcht vor HIV-Infizierung; Wissen über Infizierungsmöglichkeiten und Schutzmaßnahmen; Teilnahme an einem HIV-Antikörpertest; Intimkontakte während des letzten Jahres; Wahrscheinlichkeit eigener Infizierung.

2.2 Einstellungen von Ärzten zum Thema AIDS

AIDS beeinflußt nicht allein den Lebensbereich der Individuen, sondern betrifft darüber hinaus die Gesellschaft als Ganzes sowie gesellschaftliche Teilbereiche. So hat sich beispielsweise innerhalb des Gesundheitssystems ein weitverzweigtes Netz von Hilfsinstitutionen gebildet, die arbeitsteilig verschiedenartige Funktionen im Rahmen der AIDS-Bekämpfung wahrnehmen. Zu einem Teil haben sich dafür neue Institutionen etabliert, wie z.B. die AIDS-Hilfe, die sich in ihrer Arbeit allein auf die AIDS-Problematik konzentrieren. Darüber hinaus haben bereits bestehende Institutionen ihre Aufgabenbereiche erweitert und leisten nun vor dem Hintergrund von AIDS zusätzlich zielgruppenspezifische Beratungs- und Betreuungsarbeit (z.B. Drogenhilfe, Pro Familia etc.).

Da AIDS zunächst einmal ein medizinisch-biologisches Phänomen ist, kommt dem Arzt eine Schlüsselrolle bei der Betreuung von Betroffenen zu. Es ist zu beobachten, daß nicht nur für Ärzte im klinischen Bereich, sondern insbesondere auch für niedergelassene Ärzte die AIDS-Problematik zu einem "neuen Schwerpunkt" in ihrer Arbeit wird. Damit ergeben sich für eine steigende Zahl niedergelassener Ärzte neue Aufgaben als Berater gefährdeter und infizierter Patienten bzw. als Hausarzt in Familien mit einem betroffenen Mitglied.

Eine im Zentralarchiv archivierte Befragung von Ärzten in Köln analysiert, inwiefern sich die AIDS-Problematik auf das Arzt-Patient-Verhältnis auswirkt und ob sich in diesem Zusammenhang spezifische Einstellungen der Ärzte zu den HIV-/AIDS-Patienten und den gesundheitspolitischen Maßnahmen im Bereich von AIDS ausprägen. Die Untersuchung hat den Charakter einer Pilotstudie, da sie sich zum einen nur an diejenigen niedergelasse-

nen Fach- und Allgemeinärzte richtet, die als erste Anlaufstelle für medizinische Probleme im Zusammenhang mit AIDS prädestiniert sind. Ein in diesem Kontext entwickeltes "Konzept der ärztlichen ambulanten Primärversorgung" bezüglich AIDS[8] definiert Praktische Ärzte, Ärzte für Allgemeinmedizin, Internisten sowie Dermatologen als die Träger einer solchen Primärversorgung. Eine weitere Einschränkung erfolgte durch die regionale Begrenzung der Untersuchung auf Köln, um somit die Berücksichtigung der besonderen Bedingungen bezüglich AIDS im großstädtischen Ballungsraum zu ermöglichen.

Datensatzbeschreibung

Einstellungen zu AIDS - Ärztebefragung (Juni/Juli 1988)[9], Archiv-Nr. 1830

Primärforscher und Datenerhebung

D. Kaross, L. Fischer, Institut für Sozialpsychologie der Universität zu Köln

Grundgesamtheit und Auswahl

Untersuchungsgebiet: Köln. Postalische Befragung von niedergelassenen Ärzte in Köln, die im Bereich der ambulanten Primärversorgung bezüglich AIDS tätig sind (Praktische Ärzte, Ärzte für Allgemeinmedizin, Internisten sowie Dermatologen); 157 Befragte.

Inhalt: Erfahrungen mit HIV-Infizierten oder AIDS-Kranken in der eigenen Praxis sowie außerhalb der Sprechstunde; Anzahl der behandelten HIV-Infizierten und der durchgeführten HIV-Antikörpertests; psychische und soziale Probleme von HIV-Infizierten und AIDS-Kranken; Beweggründe der HIV-/AIDS-Patienten für den Arztbesuch; Beurteilung des HIV-Infektionsrisikos bei Ausübung des ärztlichen Berufs; Schutzmaßnahmen in der Arztpraxis; Einstellung zum HIV-Test; Einstellung zu gesundheitspolitischen Maßnahmen im Bereich der AIDS-Bekämpfung (Skala).

Ausgewählte Ergebnisse

Die umfangreichen Angaben über die Anzahl der von ihnen betreuten HIV-/AIDS-Patienten und der durchgeführten HIV-Antikörpertests läßt eine Einteilung der befragten Ärzte nach ihrem Erfahrungshorizont bezüglich der AIDS-Problematik zu[10]. Dies ermöglicht

---

8 Vgl. *Kaross, D.*\ Eine empirische Analyse von Auswirkungen der HIV-/AIDS-Problematik auf das Arzt-Patient-Verhältnis aus der Sicht von Ärzten. Universität zu Köln, Institut für Sozialpsychologie. Diplomarbeit, 15.10.1988, S. 77 ff.

9 Veröffentlichungen:
Kaross, D. (1988): a.a.O.
Kaross, D.; Knauer, R.; Fischer, L.: Probleme des Arzt-Patient-Verhältnisses im Zusammenhang mit der HIV-/AIDS-Problematik. In: *Höfling, S.; Butollo, W.* (Hrsg.): Psychologie für Menschenwürde und Lebensqualität: Aktuelle Herausforderungen und Chancen für die Zukunft. Bericht über den 15. Kongreß für Angewandte Psychologie, München, Oktober 1989: Band 1. Bonn: Deutscher Psychologen Verlag 1990, S. 135-147.

10 Als Ärzte mit viel Erfahrung (="Spezialisten") werden in dieser Studien diejenigen definiert, die im Vergleich zu ihren Kollegen überdurchschnittlich viele AIDS-Patienten betreuen oder überdurchschnittlich viele HIV-Tests durchgeführt haben. "Ohne Erfahrung" sind Ärzte, die bisher weder mit HIV-Tests

(Fortsetzung...)

eine differenzierte Betrachtung der ärztlichen Einstellungen zu ausgewählten gesundheitspolitischen Maßnahmen, die im Zusammenhang mit der HIV-/AIDS-Problematik in Fachkreisen wie auch in der Öffentlichkeit diskutiert werden.

Es ist davon auszugehen, daß sich die vom Arzt präferierte gesundheitspolitische Strategie auf seinen Umgang mit HIV-Patienten - die ja im wesentlichen die Adressaten der gesundheitspolitischen Maßnahmen darstellen - auswirkt. Ärzte sind zwar nicht immer diejenigen, die derartige Verordnungen umsetzen, ihnen käme jedoch eine grundlegende Entscheidungsmacht zu. Sie können beispielsweise mit Hilfe des HIV-Tests ermitteln, wer infiziert ist und somit einer gesundheitspolitischen Maßnahme zu unterwerfen ist.

Vor diesem Hintergrund wurde den Ärzten eine Reihe von möglichen gesundheitspolitischen Maßnahmen zur Beurteilung vorgelegt, wobei ihnen jeweils eine 5-polige Antwortskala zur Verfügung stand. Tabelle 3 verdeutlicht die Verteilung der Mittelwerte in den einzelnen Ärztegruppen für einige ausgewählte Maßnahmen.

Tabelle 3: Beurteilung von gesundheitspolitischen Maßnahmen, unterschieden nach der HIV-/AIDS-Erfahrung der Ärzte - Mittelwerte[11]

Zu bewertende Maßnahme	Bewertung durch... alle Ärzte	Spezialisten	Ärzte mit wenig Erfahrung	Ärzte ohne Erfahrung	F-Wert	p-Wert
Meldepflicht mit Namensangabe	3,28	4,05	3,31	2,71	6,18	0,003
regelmäßige serologische Untersuchungen von Risikogruppen	2,23	2,71	2,19	1,97	2,97	0,055
Verteilung von Einmal-Spritzbestecken an gefährdete Drogensüchtige	2,38	1,92	2,38	2,67	2,68	0,072
rechtliche Auflagen und Verhaltensvorschriften für HIV-Infizierte	2,61	3,08	2,58	2,36	2,56	0,081
Internierung von Aids-Kranken in speziell dafür eingerichteten Institutionen	3,85	4,33	3,78	3,71	2,50	0,085

[10](... Fortsetzung)
noch mit AIDS-Patienten konfrontiert worden sind. Die übrigen Ärzte bilden definitionsgemäß die Restkategorie mit "wenig Erfahrung". Eine detaillierte Beschreibung und Begründung dieser Unterscheidung findet sich in *Kaross et al.*, 1990, a.a.O., S. 142 ff.

[11] Die Mittelwerte beziehen sich auf die Bewertungsskala. Dabei bedeutet: 1 = sehr sinnvoll, 2 = sinnvoll, 3 = teils-teils, 4 = nicht sinnvoll, 5 = überhaupt nicht sinnvoll. Die Fragestellung lautete: "In Fachkreisen, wie auch in der Öffentlichkeit, werden unterschiedliche gesellschaftliche Strategien diskutiert, wie eine weitere Ausbreitung der Krankheit AIDS gestoppt oder zumindest verlangsamt werden kann. Dabei haben sich zwei Hauptmodelle herausgebildet: zum einen das Modell der freiwilligen Selbstkontrolle und zum anderen das Modell der amtlichen Überwachung. Für wie sinnvoll erachten Sie die hier im einzelnen aufgeführten Maßnahmen?"

2.3 ALLBUS 1988

1988 wurde erstmals AIDS als Schwerpunktthema in den ALLBUS (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften) aufgenommen, der sich an 3052 Befragte richtete.[12]

Datensatzbeschreibung

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften, ALLBUS 1988 (April 1988 bis Juli 1988)[13], Archiv-Nr. 1670

Inhalt

Kenntnis der Krankheit AIDS; wahrgenommene Berichte über AIDS in den Medien; Beurteilung des eigenen Ansteckungsrisikos bezüglich ausgewählter Kontaktformen (Skala); vermutete Entwicklung der durch AIDS verursachten Todesfälle; Beurteilung der Wahrscheinlichkeit der Entwicklung eines Medikaments gegen AIDS; Einstellung zu staatlichen Aufklärungs- und Zwangsmaßnahmen; Einstellung zu höheren Krankenversicherungsbeiträgen für HIV-Infizierte; Einstellung zur Entlassung infizierter Arbeitnehmer und einem Einreiseverbot für HIV-infizierte Ausländer; Einstellung zu einer zentralen Meldepflicht; Sorge vor einer eigenen HIV-Infektion sowie einer Infektion von Familienmitgliedern und Freunden; eigene Schutzmaßnahmen und Verhaltensänderungen; persönliches Kennen von HIV-Infizierten.

Ausgewählte Ergebnisse

Auch im ALLBUS 1988 finden sich Hinweise auf die Einstellung der Bevölkerung zu staatlichen Maßnahmen gegenüber infizierten Personen (vgl. Tabelle 4). Die namentliche Meldepflicht ist jedoch die einzige Maßnahme, zu der auch Angaben aus anderen Befragungen vorliegen (vgl. Tabellen 2, 3 und 5).

---

12 Vgl. zum ALLBUS auch die beiden Beiträge von ***Terwey*** **in** diesem Heft.

13 Veröffentlichungen:
Koch, A.: Jeder dritte junge Erwachsene ändert Sexualverhalten aus Sorge vor Ansteckung: Umfrageergebnisse zur AIDS-Prävention. In: ISI. Informationsdienst Soziale Indikatoren, Nr. 1, Januar 1989, S. 1-5.
Koch, A.: Die Krankheit AIDS: Einstellungen und Reaktionen der Bevölkerung. In: ***Müller, W.; Mohler, P.Ph.; Erbslöh, B.; Wasmer, M.*** (Hrsg.): Blickpunkt Gesellschaft: Einstellungen und Verhalten der Bundesbürger. Opladen: Westdeutscher Verlag 1990.

Tabelle 4: Einstellung zu staatlichen Maßnahmen gegenüber infizierten Personen in verschiedenen Bevölkerungsgruppen (Anteil der Befürworter in %)[14]

	N	Ausländern Einreise verweigern %	Namentliche Meldepflicht %	Höhere gesetzl. Krankenversicherung %	Arbeitnehmer entlassen %
Alter					
18-29	760	51	29	7	5
30-44	778	61	39	9	6
45-59	675	76	49	16	12
60+	800	81	60	18	13
Schulabschluß					
Hauptschule	1644	78	52	15	12
Realschule	697	63	42	12	5
Abitur	602	46	27	8	6
Wahlabsicht					
CDU/CSU	764	78	57	19	14
SPD	932	68	42	9	7
FDP	142	67	40	8	6
Grüne	191	37	17	5	3
Wahrnehmung des Ansteckungsrisikos					
sehr gering (0-4)	749	56	32	6	4
gering (5-8)	706	64	42	13	6
mittel (9-13)	634	67	44	10	10
groß (14-42)	638	83	60	21	17
Gesamt	3013	67	44	12	9

2.4 Sonstige Studien

Die Politbarometer sind monatliche Umfragen der Mannheimer Forschungsgruppe Wahlen und gehören bereits seit mehreren Jahren zu den fortlaufend in die Bestände des Zentralarchivs aufgenommenen Datensätze. Sie untersuchen im wesentlichen allgemeine politische Einstellungen der Bevölkerung sowie deren Beurteilung von Politikern auf der Basis von Stichproben mit je ca. 1000 Befragten pro Monat. Die mit Hilfe von häufig replizierten Fragen gewonnenen Zeitreihen über das politische Klima bilden darüber hinaus die

[14] Die Frageformulierung lautet: "In jüngster Zeit werden verschiedene Maßnahmen diskutiert, die gegenüber Personen ergriffen werden sollten, die mit AIDS angesteckt sind. Sagen Sie mir zu jeder Maßnahme auf dieser Liste, ob Sie sie befürworten oder ablehnen". Vgl. auch *Koch* 1990, a.a.O., S. 9.

Grundlage für die gleichnamige Fernsehberichterstattung im ZDF. Seit März 1987 wurden mehrmals Fragenkomplexe zum Thema AIDS erhoben. Vergleichbare Fragen wurden seit 1987 auch in den jährlichen Studien "Einstellungen zu aktuellen Fragen der Innenpolitik" (je ca. 2000 Befragte) durch das Institut für praxisorientierte Sozialforschung (Mannheim) gestellt, so daß damit eine umfangreiche Datenbasis für die Abbildung der Einstellungen der Bevölkerung zu gesundheitspolitischen Maßnahmen im Bereich von AIDS gebildet wird. Primärforscher der im folgenden beschriebenen Politbarometer-Studien sowie der "Einstellungen zu aktuellen Fragen der Innenpolitik" sind jeweils ***M. Berger, W. G. Gibowski, D. Roth und W. Schulte.*** Die Datenerhebung erfolgt durch MARPLAN (Offenbach) auf Basis von mehrstufig geschichteten Zufallsauswahlen von wahlberechtigten Personen in der BRD (ohne West-Berlin) mit Hilfe von mündlichen Befragungen mit standardisiertem Fragebogen.[15]

Die nachfolgende Übersicht enthält für die jeweiligen Studien eine kurze Zusammenfassung der AIDS-Fragenblöcke.

Politbarometer (März 1987), Archiv-Nr. 1642
Politbarometer (November 1987), Archiv-Nr. 1649
Politbarometer (Januar 1988), Archiv-Nr. 1687
Einstellungen zu aktuellen Fragen der Innenpolitik (Juni 1987), Archiv-Nr. 1662
Einstellungen zu aktuellen Fragen der Innenpolitik (MailJuni 1988), Archiv-Nr. 1698
Inhalt
Einschätzung der AIDS-Gefahr; Einstellung zu regelmäßigen Pflichtuntersuchungen zur Erkennung einer vorliegenden AIDS-Ansteckung; Einstellung zu einer namentlichen Meldepflicht für AIDS-Kranke durch die Ärzte; Einstellung zu einer verschlüsselten Meldepflicht; Zurückweisung von AIDS-erkrankten Ausländern

Politbarometer (Oktober 1988), Archiv-Nr. 1694
Inhalt: Wahrgenommene Bedrohung der Gesundheit der Bevölkerung durch AIDS

Ausgewählte Ergebnisse

Die mehrfache Replikation von Fragen in den Politbarometerstudien sowie in den "Einstellungen zu aktuellen Fragen der Innenpolitik" zu den möglichen staatlichen Reaktionen auf AIDS ermöglicht die Abbildung des Meinungsklimas in den Jahren 1987 und 1988 (vgl. dazu die Tabelle 5 auf der folgenden Seite).

---

[15] Das Politbarometer im Oktober 1988 wurde allerdings mit Hilfe einer telefonischen Befragung (CATI) erhoben.

Tabelle 5: Einstellung zu staatlichen Maßnahmen gegenüber HIV-Infizierten und AIDS-Kranken (Angaben in Prozent)

	Politbarometer "Einstellungen..." März 1987 N=1029	Juni 1987 N=2045	Politbarometer November 1987 N=1007	Politbarometer Januar 1988 N=1034	"Einstellungen..." Mai/Juni 1988 N=2078
Frage 1[16] *regelmäßige HIV-Tests für die Bevölkerung*					
dafür	57.0	47.8	49.5	47.4	45.8
dagegen	42.9	51.4	50.5	52.4	54.0
Frage 2a[17] *namentliche Meldepflicht*					
dafür	50.1	47.1	43.8	45.0	48.5
dagegen	49.9	52.4	56.2	54.9	51.3
Frage 2b[18]					
Nur an diejenigen Befragten gerichtet, die sich gegen eine namentliche Meldepflicht aussprachen: Einstellung zur verschlüsselten Meldepflicht:					
dafür	60.4	57.4	57.1	nicht	69.0
dagegen	39.6	42.6	42.9	erhoben	31.0
	(n=513)	(n=1064)	(n=564)		(n=1062)

16 Der genaue Fragentext lautet "Es wird gefordert, daß sich die ganze Bevölkerung in regelmäßigen Abständen auf AIDS-Ansteckung hin untersuchen lassen muß. Sind Sie für eine solche Pflichtuntersuchung oder sind Sie dagegen?"; (Politbarometer März 1987 und November 1987)
Der Wortlaut in den Befragungen der "Einstellungen zu aktuellen Fragen der Innenpolitik" wurde nur unwesentlich modifiziert. Im Politbarometer vom Januar 1988 wurde eine vollkommen neue Frageformulierung gewählt, die inhaltlich jedoch vergleichbar bleibt.
In weniger als 1% der Fälle wurde die Antwort verweigert.

17 Der genaue Fragentext lautet: "Es wird auch gefordert, daß die Ärzte jeden mit Namen melden, bei dem der AIDS-Erreger festgestellt wird. Sind Sie für eine solche namentliche Meldepflicht oder sind Sie dagegen?" (Politbarometer im März 1987 und November 1987).
Der Wortlaut in den Befragungen der "Einstellungen zu aktuellen Fragen der Innenpolitik" wurde nur unwesentlich modifiziert. Im Politbarometer vom Januar 1988 wurde eine vollkommen neue Frageformulierung gewählt, die inhaltlich jedoch vergleichbar bleibt.
In weniger als 1% der Fälle wurde die Antwort verweigert.

18 Diejenigen Befragten, die eine namentliche Meldepflicht ablehneten, wurden zusätzlich gefragt: "Es ist geplant, daß alle Personen, bei denen der AIDS-Erreger festgestellt wird, ohne Nennung ihres Namens mit einer Kennziffer gemeldet werden müssen. Sind Sie für eine solche verschlüsselte Meldepflicht oder sind Sie dagegen?" (Politbarometer im März 1987 und November 1987).
Auch zu dieser Fragestellungen wurden im Politbarometer (Januar 1988) sowie in den Untersuchungen zu den "Einstellungen zu aktuellen Fragen der Innenpolitik" leicht modifizierte Formulierungen gewählt.

In der Tabelle 5 lassen sich für den Zeitraum von März 1987 bis Juni 1988 einige sehr deutliche Veränderungen in den Einstellungen der Bevölkerung zu gesundheitspolitischen Maßnahmen zur AIDS-Bekämpfung ablesen. Die Akzeptanz bezüglich regelmäßigen serologischen Pflichtuntersuchungen ist beispielsweise um mehr als 10 Prozentpunkte gesunken. Die im März 1987 noch leicht überwiegende Zustimmung zur namentlichen Meldepflicht hat sich seit Juni 1988 in eine überwiegende Ablehnung gewandelt, die jedoch seit 1988 wieder mit steigenden Akzeptanzwerten verbunden ist.

Es stellt sich die Frage, wie sich diese plötzlichen und zum Teil sehr gravierenden Wandlungen im Meinungsklima erklären lassen. Ohne den Anspruch einer fundierten Analyse zu erheben, sollen an dieser Stelle einige Überlegungen angestellt werden, um zu ersten Erklärungsansätzen zu gelangen.[19]

Dazu bietet sich eine genauere Betrachtung des gesellschaftspolitischen Umfeldes zu den Befragungszeitpunkten an. In den Jahren 1986 bis 1988 wurde das Thema AIDS in der Öffentlichkeit sehr ausgiebig erörtert. Nahezu täglich fanden sich Schlagzeilen in der Tagespresse, die jedoch in vielen Fällen auch zur weiteren Verunsicherung in der Bevölkerung beitrugen. Gleichzeitig wurde von der BZgA versucht, mit bundesweit angelegten Informations- und Aufklärungskampagnen bestehende Ängste und Unkenntnisse zu beseitigen. Einen umfangreichen Raum nahm die Diskussion um mögliche Präventionsmaßnahmen ein, die das einzige Mittel darstellen, um die Verbreitung der tödlichen Krankheit und die Zahl der Opfer wirksam zu vermindern, solange es nicht gelingt, einen Impfstoff gegen HIV oder wirksame Therapeutika gegen AIDS zu entwickeln. In diesem Zusammenhang wurde insbesondere über die angemessene gesundheitspolitische Strategie des Staates nachgedacht.[20] 1987 wurde von der Bayerischen Staatsregierung ein "Maßnahmenkatalog zur Verhütung und Bekämpfung der Immunschwächekrankheit AIDS" beschlossen.[21] In den Medien erfolgte eine vereinfachte Polarisierung zwischen dem "Modell der amtlichen Überwachung, verbunden mit staatlicher Regulierung" und dem "Modell der Aufklärung und freiwilligen Selbstkontrolle", die nahezu untrennbar mit den Namen *Gauweiler* auf der

---

19 Auf die Tatsache, daß die Frageformulierungen in den verschiedenen Untersuchungen nicht vollkommen identisch sind und somit möglicherweise einen Anteil der hier abgebildeten Meinungsvarianz verursachen, wurde bereits hingewiesen und soll an dieser Stelle nicht weiter vertieft werden.

20 Vgl. z.B. *Frösner, G:*. Wie kann die weitere Ausbreitung von AIDS verlangsamt werden? In: AIDS-Forschung (AIFO), 2. Jg., Heft 2, 1987; *Rosenbrock, R.:* AIDS kann schneller besiegt werden. Gesundheitspolitik am Beispiel einer Infektionskrankheit. Hamburg: VSA-Verlag 1987 sowie *Süssmuth, R.:* AIDS. Wege aus der Angst. Hamburg: Hoffmann und Campe 1987.

21 Vgl. zum Bayerischen AIDS-Präventionskonzept: *Satzinger, W.; Bujok, E.:* Zwischen Seuchenprävention und Sozialfürsorge. Bemerkungen über AIDS-Beratung unter bayerischen Bedingungen. In: *Rosenbrock, R.; Salmen, A.* (Hg.): AIDS-Prävention. Berlin: Ed. Sigma Bohn 1990, S. 67-76.

einen sowie ***Süssmuth*** auf der anderen Seite verknüpft wurden.[22]

Diese kontrovers und zum Teil emotional geführte Diskussion um AIDS und die gesellschaftlichen Konsequenzen führte zu den erwähnten Verunsicherungen, die für die hier ermittelten Schwankungen im Meinungsklima ursächlich sein könnten.

3. Zusammenfassung

Ziel dieses Beitrags war es, Einblick in einige empirische Daten zum Thema AIDS zu geben. Neben Datensatzbeschreibungen wurde eine Reihe von Ergebnissen ausgewählt, die Anhaltspunkte für die gesellschaftlichen Implikationen von AIDS aufzeigen. Bisher mangelt es zwar an langfristig angelegten Wiederholungsbefragungen, dennoch konnten anhand verschiedener Umfragen mit vergleichbaren Frageformulierungen erste Bewegungen im gesellschaftlichen Meinungsklima festgestellt werden. Insbesondere die Einstellungen zu einigen zeitweise diskutierten staatlichen Maßnahmen zur Bekämpfung von AIDS, wie z.B. die namentliche Meldepflicht, lassen einige zum Teil gravierende Schwankungen im Verlauf eines Jahres erkennen. Erkennbar wird auch, daß die gesundheitspolitische Diskussion unter ärztlichen Experten unter anderen Gesichtspunkten geführt wird und daß sich deren Einstellungen von denen in der Bevölkerung zum Teil unterscheiden.

Bisher ist der zeitliche Horizont zu kurz, um langfristige Trends erkennen zu können. Erst wenn exakte Replikationen der Fragenprogramme über einen weiten Zeitraum vorliegen, werden fundierte Aussagen über Meinungsänderungen möglich sein.

Dietmar Kaross

---

[22] Vgl. z.B. *Bohn, R.:* Droht die Gauweilerisierung der Republik? In: Forum Wissenschaft, 1987, Heft 2, S. 11-15.
Einen allgemeinen Überblick über die massenmediale Behandlung des Themas AIDS gibt *Mackenthun, G.:* Massenmedien und AIDS-Berichterstattung. In: *Rosenbrock, R.; Salinen, A.* (Hg.): AIDS-Prävention. Berlin: Ed. Sigma Bohn 1990, S.77-83.

Studien aus der ehemaligen DDR im Zentralarchiv

Seitdem im Rahmen der Erhebung der sogenannten USIA-Studien Ende der 50er Jahre Bürger aus der DDR bei Berlin-Besuchen zur Wiedervereinigung befragt wurden, sind keine weiteren Datensätze aus der DDR verfügbar gewesen. Vor wenigen Wochen erhielten wir sieben Datensätze von Erhebungen aus den 80er Jahren. Als Dokumente der Markt- und Meinungsforschung in der ehemaligen DDR haben wir die Datensätze in unsere Bestände aufgenommen. Ein aktueller Datensatz aus dem Frühjahr 1990 "Leben in der DDR" ist als erster Zugang aus der Zeit nach dem Fall der Mauer hervorzuheben. Weitere Datensätze werden folgen. Alle Daten liegen als SPSS-Systemfiles vor. Wir stellen die uns übermittelten Beschreibungen dieser Datensätze vor.

ZA-Archiv-Nr. 1866 und 1867
Zur sozialen Entwicklung im Konsumgüterbinnenhandel der DDR 1982 und 1988
Primärforscher und Datenerhebung
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Arbeits- und Lebensbedingungen im Dienstleistungssektor unter besonderer Berücksichtigung des Konsumgüterbinnenhandels, des weiteren Leistungsverhaltens, Leistungsbewertung, Arbeitskollektiventwicklung, Wettbewerb, Leistungstätigkeit.
Untersuchungsgebiet: Berlin, Schwerin, Karl-Marx-Stadt.
Auswahl von Funktionsgruppen im Einzelhandel (Konsum und HO) und im Großhandel (Waren täglicher Bedarf).
Befragtenzahl: 1662 bzw. 2610; Variablen: 317 bzw. 376.

ZA-Archiv-Nr. 1868 und 1869
Kinderwunsch 1982 und 1987
Primärforscher und Datenerhebung
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Objektive und subjektive Bedingungen für die Geburt von Kindern (Wertestrukturen, Bedürfnisstrukturen), Stärke und Inhalt des Kinderwunsches.
Untersuchungsgebiet: Berlin, Wismar, Riesa, Gemeinden des Kreises Templin.
1982: Auswahl von Frauen im Alter von 18 bis 40 Jahren.
1987: Auswahl von Männern (33%) und Frauen im Alter von 18 bis 40 Jahren
Befragtenzahl: 2488 bzw. 4136; Variablen: 221 bzw. 265.

ZA-Archiv-Nr. 1870
Leistungsverhalten und Aktivität bei der Durchsetzung des wissenschaftlich-technischen Fortschritts (1983)
Primärforscher und Datenerhebung
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Leistungsverhalten in Innovationsprozessen. Arbeits- und Lebensbedingungen, Leistungsvoraussetzungen und -bewertung, Arbeitskollektivbeziehungen, Leitungstätigkeit, Wertorientierungen.
Untersuchungsgebiet: Berlin, Potsdam, Frankfurt/Oder. Auswahl von Arbeitsgruppen aus Industriebetrieben und Instituten der Akademie der Wissenschaften der DDR.
Befragtenzahl: 1097; Variablen: 570.

ZA-Archiv-Nr: 1871
Friedensengagement und Leistungsverhalten sozialistischer Produzenten (Nov. 1987)
Datenerhebung
Zentralinstitut, Universität Halle/Wittenberg;
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Ziel war die Analyse eines neuen Herangehens der Werktätigen der DDR an weltpolitische Fragen. Erforscht wurden Bewußtseinsinhalte, politisches Wissen und Denken. Handlungsbereitschaft für friedensbewußtes Verhalten.
Untersuchungsgebiet: Bezirk Frankfurt/Oder. Auswahl von Männern und Frauen (37%) aus 8 Betrieben.
Befragtenzahl: 1524; Variablen: 245.

ZA-Archiv-Nr. 1872
Frauen in höheren Leitungsfunktionen im Bereich des Schwermaschinen- und Anlagenbaus (April 1989 bis August 1989)
Primärforscher und Datenerhebung
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Im Ergebnis der empirischen Erhebung sollten Ursachen für den geringen Anteil an Frauen in höheren Leitungsfunktionen aufgedeckt werden. Hemmende Faktoren für die Bereitschaft zur Übernahme von Leitungstätigkeit bei Hochschulabsolventinnen sollten erforscht werden, um Veränderungen in diesem Bereich herbeiführen zu können.
Untersuchungsgebiet: DDR. Auswahl von 454 Männern und 777 Frauen in leitenden bzw. gehobenen Positionen in 16 Betrieben aus dem Bereich Schwerindustrie und Anlagenbau.
Befragtenzahl: 1231; Variablen: 412.

ZA-Archiv-Nr. 1873
Leben in der DDR (Januar 1990)
Primärforscher und Datenerhebung
Michael ***Häder,***
Institut f. Soziologie und Sozialpolitik an der Akademie der Wissenschaften der DDR, Berlin.
Inhalt
Denken und Handeln der Menschen in der DDR während eines grundlegenden politischen und ökonomischen Systemwandels.
Hauptgebiete der Studie sind: Umwelt, Demokratie, Arbeit, Familie, Freizeit, Dienstleistungen, Wohnen. Anhand eines Handlungsmodells wurden Wertorientierungen (Lebens-) Zufriedenheit, Informiertheit, soziale Erwartungen (Normen), Zukunftserwartungen und Handlungsbereitschaft erhoben.
Untersuchungsgebiet: DDR, Bezirke Potsdam, Leipzig und Schwerin (12 Orte). Es war beabsichtigt alle Personen ab 18 Jahren in Haushalten zu befragen.
Befragtenzahl: 1528 Personen aus 828 Haushalten; Variablen: 370.

Kreisdaten der 227 ehemaligen DDR-Kreise

Die Forschungsbereiche und Forschungsgesellschaften der infas-Gruppe erstellen seit 1959 mit den Instrumenten der empirischen Wirtschafts- und Sozialforschung Planungsunterlagen und Entscheidungshilfen für Wirtschaft, Wissenschaft, Verwaltung, Politik und Medien. Das **infas**-Institut für Absatzplanung in München informiert, daß nun maschinenlesbare Aggregatdaten auf Kreisebene für die Region der ehemaligen DDR vorliegen.
Aus dem Inhalt:
Für jeden der 227 Kreise der DDR (mit Zwischensummen für die 15 Bezirke und Summe DDR insgesamt) enthalten die Tabellen die wichtigsten verfügbaren Daten, wie
-Wohnbevölkerung nach Geschlecht und Alter
-Haushalte (nach Größenklassen)
-Geburten, Sterbefälle, Eheschließungen
-Sparguthaben der Bevölkerung
-Berufstätigkeit (nach Wirtschaftszweigen)
-Industrieproduktion
-Landwirtschaft (Flächennutzung/Viehbestand)
-Kfz-Bestand
-Einzelhandelsumsatz
-Bautätigkeit
-Wohnungsbestand
-Gesundheitswesen
sowie ein Kurzprofil jedes einzelnen Kreises mit Erläuterungen zu den wesentlichen Indikatoren in Textform. Die Kreisdaten sind auf Diskette und/oder als Ausdruck (Papier) über **infas**-Absatzplanung, Hildegardstr. 2, 8000 *München* 22 zu beziehen. Anfragen können auch an das Zentralarchiv gerichtet werden.

Archivexperten diskutieren Datenbankprobleme

Das dritte CESSDA-Experten-Seminar fand vom 27. bis 30. 9. 1990 im Zentralarchiv zum Thema "Advanced Techniques of Data Management and Documentation" statt. Mitarbeiter aus neun europäischen Datenarchiven versammelten sich zum Austausch von Informationen, Erfahrungen und Perspektiven. Vertreten waren die Archive aus Dänemark, Großbritannien, Norwegen, Frankreich, Österreich, Schweden, den Niederlanden, Ungarn und Deutschland; d.h. neben den bereits seit Jahren etablierten Einrichtungen waren auch die "jüngeren" Institute beteiligt - eine durchaus gute Voraussetzung für den Austausch von Erfahrungen.

Es ist zu beobachten, daß die Zahl der Datensätze in der Sozialforschung steigt, die nicht mehr im "flat file"- Format abzubilden sind. Als "flat file"- Format werden die in der Sozialforschung am häufigsten vorkommenden Datenstrukturen bezeichnet: jeder Untersuchungseinheit (z.B. jedem/jeder Befragten) ist die gleiche Anzahl von Variablen zugeordnet. Ein solcher Datensatz kann problemlos in einer zweidimensionalen (rechteckigen) Matrix dargestellt und als solche auch von den meisten Statistiksoftwarepaketen ausgewertet werden.

Der erste Teil des Seminars war den sog. "komplexen" Datensätzen gewidmet, also jenen Datenstrukturen, die über das "flat file"- Format hinausgehen. Das sollte sowohl für die Aufbereitung der Daten selbst gelten als auch für die Erstellung von Codebüchern und Dokumentationen zu diesen Daten. Zunächst einmal wurde versucht zu präzisieren, was unter "Komplexität" zu verstehen ist: sind komplexe Dateien einfach nur "große" Dateien, die wegen ihrer Größe unter technisch erschwerten Bedingungen zu bearbeiten sind, sind es hierarchisch gegliederte Strukturen oder als Netzwerke organisierte Files, oder gibt es etwa einen "Komplexitäts-Komplex" (psychologischer Art)?

Für die Archive stellt sich die Frage, ob der Aufwand zur Erstellung einer "komplexen" Datenbank vertretbar ist, zumal die Mehrzahl der Sekundäranalysen dann doch wieder mit "rechteckigen" Dateien vorgenommen wird. Aus der Archivperspektive ist es vor allem eine Frage der (Personal-)Kapazität, ob komplexe Datensätze als solche verarbeitet, dokumentiert und für den Zugriff bereitgehalten werden. Trotz seiner wechselvollen Geschichte wurde SIR, als Software gerade für den sozialwissenschaftlichen Bereich, wieder positiv hervorgehoben. Gleichzeitig wurde aber eine grundsätzliche Schwäche dieser wie auch anderer spezieller Datenbanksoftware herausgestellt: deren Sprache.

SQL ist zwar als Standardsprache in vielen DBMS implementiert, es gibt aber nahezu ebensoviele SQL -"Dialekte" wie DBMS-Hersteller. Zudem erfordert ein kompetenter Umgang mit den Systemen eine intensive und zeitaufwendige Beschäftigung mit der Datenbanksprache -

dies kann aber nicht von jedem Nutzer eines Archivs erwartet werden, der ja in erster Linie an seiner Analysesoftware interessiert sein muß. Im Rahmen dieser Diskussion wurde die Kumulation der ALLBUS-Studien als ein Beispiel für die Aufbereitung und Dokumentation eines komplexen Datensatzes unter Verzicht auf ein kompliziertes Datenbanksystem vorgestellt. Dabei wird sowohl für die Datenbearbeitung als auch für die Verwaltung der Meta-Information SPSSx verwendet. Das Programmsystem SPSSx hat den Vorteil, daß es einerseits genügend Flexibilität für die verschiedenen Aufgaben der Kumulation hat, daß es weiterhin mit relativ geringem Speicherplatz für komprimierte Systemfiles auskommt und daß es andererseits nach wie vor das in der Profession am weitesten verbreitete Analysepaket ist. Die Aufbereitung der ALLBUS-Studien inklusive deren Kumulation wird im Zentralarchiv von einer eigenen ALLBUS-Arbeitsgruppe geleistet. Also gilt auch hier das Argument, daß "komplexe" Aufgaben in der Regel in den Archiven nicht ohne eigene (zusätzliche) (Personal-Kapazität erbracht werden können. Die Quintessenz der Diskussionen der internationalen Archiverfahrungen mit "komplexen" Datensätzen lautete: "keep it simple as long as possible".

Ein weiterer Schwerpunkt des Seminars war die Rückgewinnung von Information. In jedem der beteiligten Archive wird damit natürlich ein essentielles Thema angesprochen. Die Basisfragen lauten: Wie bediene ich eine Benutzeranfrage? Welche Quellen dienen mir selbst zur Information? Welche Quellen kann ich externen Interessenten zur Verfügung stellen? Ein Verzeichnis der gespeicherten Datensätze auf Papier gibt es in jedem Datenarchiv. Globale Informationen sind dadurch relativ schnell zu übermitteln, bzw. der Benutzer kann diese Quelle auch selbst nutzen. Das Ergebnis einer solchen Recherche lautet zumeist: "Ich möchte den Datensatz und die Dokumentation der Studie XYZ beziehen."

Komplizierter ist die Art der Anfrage: "Haben Sie irgendetwas zum Thema:...?" Hier sind in der Regel computergestützte Retrievalsysteme gefragt, am nützlichsten sind darunter solche, die in der Lage sind, bis auf die Indikatorebene hinabzugehen. Aber genau dort beginnen die Probleme und dort scheiden sich auch die Geister. Ist ein Retrievalsystem schwerpunktmäßig ein "in-house"-Instrument oder soll es (via Computernetzwerk) der Klientel angeboten werden können? Entsprechend stellt sich die Frage, ob ein solches System auf einem PC unter MS DOS laufen soll, auf einer Workstation mit Netzwerkanbindung oder auf einer Mainframe mit erleichtertem Fernzugriff. Daran anknüpfend wiederum ist die Frage der Benutzeroberfläche von Interesse. Für den PC-Bereich gibt es z.B. eine Menge von Retrievalsystemen, die fortgeschrittene Suchmöglichkeiten eröffnen - diese sind jedoch für den Fernzugriff nur wenig geeignet und in ihrer Performance von den verfügbaren Prozessoren abhängig; andere Systeme werden für den Mainframe-Bereich angeboten, sind aber bei den speziellen Informations-Rückgewinnungsfunktionen nur oberflächlich ausgestattet - bzw. zumeist sehr kostspielig. So wird die Retrieval-Frage zu einer prinzipiellen Frage und von den Archiven unterschiedlich beurteilt. Herauszustellen ist jedoch, daß keine Ideallösung in

Sicht ist, daß sich die beteiligten Institute mehr oder weniger mit dem Status quo arrangieren, jedoch die jeweils realisierte Lösung nicht als endgültig ansehen. In der Tat bewegt sich ja auf diesem Gebiet sowohl in der Soft- als auch in der Hardwareentwicklung einiges, so daß eine (nahezu) Ideallösung nach Meinung vieler für die Zukunft nicht ausgeschlossen wird. Diese böte dann auch die Chance, sich zumindest europaweit auf einen Standard zwischen den Archiven zu einigen.

In der abschließenden Diskussion wurde eine informelle Plattform ins Leben gerufen, auf der sich die "CESSDA-Experten" in unregelmäßigen Abständen über archiv-alltägliche Probleme, Wünsche oder auch über Lösungen informieren. Sinn dieses Austausches in der Zeit zwischen den Expertenseminaren soll vor allem sein, Doppelentwicklungen zu vermeiden und, wenn möglich auch in den nachgeordneten Bereichen, Standardisierungen zu erreichen.

Die Institution der CESSDA-Experten-Seminare hat sich auch diesmal als hilfreich und gewinnbringend erwiesen. Das manifestiert sich nicht zuletzt an der großen Teilnehmerzahl (22 Personen) dieser Veranstaltung. In der "Manöverkritik" wurde dieser Faktor jedoch von einigen Teilnehmern relativiert und es wurde der Wunsch geäußert, in Zukunft die Themenstellung mehr einzuschränken, damit mehr in Einzelheiten diskutiert und demonstriert werden kann. Dieser Wunsch impliziert jedoch, daß die Anzahl der Experten, die an einem solchen Seminar teilnehmen, geringer wird, weil z.B. die "kleineren" Archive durch mangelnde personelle oder materielle Ausstattung eine weitgehende Spezialisierung nicht leisten können und dadurch benachteiligt werden. Doch gerade die Integration dieser Archive in das internationale Netzwerk der Serviceeinrichtungen für die Sozialforschung sollte auch ein Ziel der Expertenseminare sein und bleiben.

Rolf Uher

Überlegungen zur Konzeption des Daten- und Informationsmanagements für den ALLBUS

von Michael Terwey

1. Kurzevaluation von Datenmanagementsystemen unter dem Gesichtspunkt der ALLBUS-Anwendung

Die auf S. 18-29 dieses Heftes vorgestellten ALLBUS-Daten werden vom Zentralarchiv nicht nur als Einzelstudien aufbereitet, sondern auch insgesamt in einen gemeinsamen **Datenpool** integriert. Aus diesem Gesamtpool werden gemäß der für die Benutzerversion des **kumulierten ALLBUS** entwickelten Konzeption, Teilmengen extrahiert und an die Anwender weitergegeben. Die technischen Probleme bei der Zusammenfassung von Daten und Dokumentationen über die Daten wachsen im Zeitverlauf mit der steigenden Anzahl der zu berücksichtigenden Einzelstudien.

Für die umfangreichen und komplex strukturierten Daten und Dokumentationen zum ALLBUS wurden deshalb im Zentralarchiv verschiedene Alternativen eines universelleren "Data Base Management Systems" (DBMS) überprüft. Neben anderen DBMS-Anwendungen bestehen im Zentralarchiv bereits mehrere SIR-Datenbanken (z. B. "Zeitbudget-Studie" (vgl. ***Wilke/Römer*** 1981), "Pilotstudie Bildschirmtext", "Reiseanalysen" und im Zentrum für Historische Sozialforschung (ZHSF) "Abgeordnete Deutscher Nationalparlamente, 1848-1933" (vgl. *Ponemereo* 1988) sowie "Die Bevölkerung der Exulantenstadt Glückstadt, 1616-1652"). Diese SIR-Datenbanken sind zum Teil in der Anwendung unterschiedlich effizient. Während beispielsweise die Reiseanalysen als vollständige, rechteckige Dateien abgespeichert wurden, legt die Abgeordneten-Datenbank aufgrund der unterschiedlichen Datentypen (Aggregatdaten, Individualdaten, wiederholte Beobachtungen, bestimmte Beobachtungen liegen nur für Abgeordnete aus bestimmten Regionen vor) ein differenzierteres Abspeicherungsschema nahe, das in einer komplexen hierarchischen Struktur realisiert wurde. Der Effizienzgewinn ist demzufolge bei der Abgeordneten-Datenbank beträchtlich, während die Reiseanalysen-Datenbank weniger Verbesserungen gegenüber herkömmlichen Datenabspeicherungen zeigt. Letzteres ist in der "normalen" Datenstruktur der Reiseanalysen bedingt, die zur einfachen Abspeicherung von rechteckigen Dateien (flat files) führte, so daß der Einsatz von SIR weniger Gewinn bringen mußte. Die Rechenzeiten für einfache Auswertungen liegen bei den Reiseanalysen höher als bei vergleichbaren Auswertungen mit SPSS. Andererseits erleichtert hier der Einsatz von SIR die Generierung von Zeitreihen aus den einzelnen "flat files" und Konkordanzübersichten.

Angesichts der weiteren Entwicklung von Datenbanksystemen ist in jüngster Zeit die Benutzerfreundlichkeit und die Performanz von SIR gelegentlich kritisiert worden. Von seiner ursprünglichen Konzeption her ist SIR eigentlich ein hierarchisches DBMS (vgl. *Schlageter/ Stucky* 1983: 71-80). Einer der Vorteile des SIR-Systems liegt darin, daß seine Kommandosprache und seine Fazilitäten von den uns bekannten DBMS noch am ehesten mit denen des SPSS vergleichbar sind. Für Zugriffe auf die Daten, die sich an der vorgegebenen Hierarchie in der Datenbankstruktur orientieren, ist SIR zudem ein vergleichsweise schnelles System. Die neueren **relationalen** Datenbanken (RDBMS, vgl. *Schlageter/Stucky* 1983: 80-90) sind dagegen in ihrer Struktur und Anwendung flexibler, weshalb SIR von einigen Anwendern bereits als überholt angesehen wird. Dennoch konnten bei einem Vergleich von DBMS-Anwendungen auf dem ZA-Datenbankseminar 28.-31.03.89 (vgl. *Mochmann* 1988; *Schlögell* 1989) auch die vorgestellten SIR-Anwendungen in vieler Hinsicht überzeugen.

Auf dem 3. CESSDA-Expertenseminar (Committee of European Social Science Data Archives) fand ferner vom 27.-30.09.90 im Zentralarchiv ein Erfahrungsaustausch über die in den europäischen Datenarchiven DBMS-mäßig vorliegenden Datensätze und die hierfür verwendete Software statt (vgl. auch den Beitrag von Rolf *Uher* in diesem Heft). Soweit dort Datensätze wegen ihrer Komplexität bisher in einem DBMS gehalten werden, findet vorwiegend SIR Verwendung. Hervorgehoben wurde die Fähigkeit von SIR, in der Datenbank auch ähnlich wie SPSS "labels" zu halten und für die Studienweitergabe aus der Datenbank SPSS-Setup und SPSS-Datensatz extrahieren zu können. Die SPSS-gesteuerte, automatisierte Erstellung eines Setups aus einem SPSS-Systemfile ist seit der Einführung von SPSSx nicht unproblematisch (siehe auch Abschnitt II).

Als alternative DBMS zu SIR bieten sich insbesondere die relationalen Systeme SQL/DS, DB2, ORACLE und INGRES an. Durch den exemplarischen Einsatz einiger dieser Systeme wurden im Zentralarchiv bereits weitere Erfahrungen gesammelt:

Für die deutschen Wahlstudien aus den Jahren 1953-1987 wurde eine INGRES-PC-Datenbank im Zentralarchiv entwickelt (vgl. ***Uher*** 1989,1990). Diese Datenbank enthält nicht die Daten dieser Studien selbst, sondern Informationen über die Studien, mit Hilfe derer die Benutzer über alle Wahlstudien hinweg relevante Indikatoren für eine bestimmte Thematik ermitteln können.

Die eingangs genannte SIR-Datenbank "Die Bevölkerung der Exulantenstadt Glückstadt, 1616-1652" wurde von Ralph ***Ponemereo*** für Vergleichszwecke auch in INGRES-PC angelegt. Beide Formen dieser Datenbank sind im Zentralarchiv unter der Studiennummer S8002 erhältlich (vgl. Zentralarchiv für empirische Sozialforschung 1990: 668).

Im Rahmen des bereits erwähnten ZA-Datenbankseminars wurde die Verwendung der INGRES-Mainframe-Version mit einer ISSP-Testdatenbank von Thomas ***Flory*** demonstriert. Als weitere RDBMS-Testanwendungen im Zentralarchiv sind eine SQL/DS-Datenbank mit

Bandarchivdaten und eine ORACLE-PC-Datenbank mit einem kleinem Datensubset aus dem kumulierten ALLBUS zu nennen, die beide von Matthias *Fischer* generiert wurden. Grundsätzliche Schwierigkeiten ergaben sich bei diesen Anwendungen weniger bei der Verwaltung numerischer Daten, sondern bei längeren Textfeldern, wie sie etwa in Codebüchern enthalten sind (vgl. auch ***Bender/Bien*** 1988: 117). Letztere Aufgabenstellung wurde exemplarisch von Lorenz *Gräf* in dem dBASE-Clipper-System FRAVIS gelöst (vgl. *Gräf* 1988).

Eine weitere Problematik auch neuerer DBMS ist, daß sie nur teilweise mit einfachen Übergabemöglichkeiten an verbreitete Analysepakete wie SPSS, BMDP oder SAS versehen sind, da sie vorwiegend für Anwendungen außerhalb der Sozialwissenschaften konzipiert sind. Beispiele für bereits vorliegende Ausnahmen sind das SPSS-Interface zu ORACLE (SPIRIT*SQL) und die Schnittstelle, welche SPSSx für SQL/DS und DB2 als Zusatzprogramm (SPSSx Capture) anbietet.[1]

Ein gravierender Nachteil von primär für kommerzielle Anwendungen entwickelten Datenbankprogrammen ist, daß viele dieser Systeme an sich keine definierten fehlenden Werte kennen. In sozialwissenschaftlichen Daten sind fehlende Werte jedoch oft unvermeidlich und als solche sorgfältig zu behandeln. SIR kann neben einem "system missing value" wenigstens drei unterschiedlich definierte fehlende Werte aufnehmen, was allerdings für einige ALLBUS-Variablen noch zu wenig ist. Auch ORACLE hat die Möglichkeit, fehlende Angaben als solche gänzlich aus den physikalisch abgespeicherten Daten auszuschließen (vgl. ***Bender/Bien*** 1988: 116). Andere Typen von "missing value" sind meines Wissens in den genannten RDBMS nur durch ad hoc-Programmierung zu behandeln.

Insgesamt ist gerade in der gegenwärtigen Phase noch nicht genau absehbar, welches DBMS für die Sozialwissenschaften längerfristig optimal sein wird. Diese Feststellung gilt um so mehr als auch die Anwender relationaler Datenbanken noch in vieler Hinsicht mit den realisierten Standards unzufrieden sind und an Verbesserungen, etwa der Standardabfragesprache **SQL** ("Structured Query Language"; vgl. *Archilles* 1989, *Eilers* et al. 1988), gearbeitet wird. Obgleich SQL die Funktion einer standardisierten Abfragesprache wahrnehmen soll, sind bei komplexeren Abfragen Besonderheiten der jeweiligen DBMS in der Syntax zu beachten, so daß nur bedingt von Standardisierung und von systemübergreifender Kompatibilität gesprochen werden kann (vgl. auch *Bender/Bien* 1988: 115).[2] Dies tritt nach unserer Erfahrungen

---

1 SPIRIT*SQL wird von AGRIMATHICA WAGENINGEN B.V. für PC's unter MS-DOS und für VAX angeboten. SPSSx-Capture kann in der IBM-Welt eingesetzt werden, für die speziell DB2 und SQL-DS entwickelt wurden. Weitere Verbindungsmöglichkeiten bestehen zwischen SAS und ORACLE sowie zwischen SIR und SPSS oder BMDP.

2 Am Rande sei darauf hingewiesen, daß es auch eine Fülle von Verbesserungs- und Ergänzungsvorschlägen für die der strategischen Konzeption einer Datenbank zugrundezulegenden Datenmodelle gibt. Stellvertretend sei hier nur auf die "Entity-Relationship"-Modelle hingewiesen (vgl. ***Chen*** 1976,1980, ***Mayr et al.*** 1987, ***Abbenhardt 1990).***

insbesondere bei den "string"-Funktionen zutage, mit deren Hilfe auch Textfelder verarbeitet werden können.

Für die Abspeicherung der ALLBUS-Daten im Zentralarchiv ergab sich unter diesen Voraussetzungen die Frage, welche Form der Datenhaltung unter den gegebenen Voraussetzungen die wichtigeren Vorteile bietet. Eine vergleichende Untersuchung von Datenmanagementfunktionen mit SIR und SAS *(Althainz* 1987) ergab zwei Hauptunterschiede. SAS benötigte für Paneldaten der Lebensverlaufsstudie des Sonderforschungsbereichs 3 der DFG 25% mehr Speicherplatz als SIR und brauchte zum Erzeugen eines für komplexere Panelanalysen geeigneten Datensatzes um 30% längere Rechenzeit. Angemerkt werden muß hierzu, daß die Daten der Lebensverlaufsstudie auch im SAS als "Pseudodatenbank" in Systemdateien abgespeichert wurden. Eingelesen wurde in die SAS-Datenbank ein Rohdatenextrakt aus dem SIR. Die SAS-Dateien beanspruchten annähernd doppelt so viel Speicherplatz wie die Rohdaten *(Althainz* 1987: 69). Der Vergleich zwischen SIR und SAS dürfte in dieser Hinsicht nur eingeschränkt auf Erfahrungen mit dem SPSSx auszudehnen sein (s.u.).[3]

Die Gegenüberstellung von SIR und SAS gilt jedoch auch weiterhin nur eingeschränkt, wenn es sich im Unterschied zur Lebensverlaufsstudie nicht um Paneldaten handelt. Paneldaten, d. h. wiederholte Beobachtungen für gleiche Fälle, gehören zu den Datentypen, bei denen der Einsatz eines komplexeren DBMS einen größeren Performanzgewinn ermöglicht. Auch bringt das Überspielen von Daten in eine SIR-Datenbank weniger Speicherplatzvorteile, wenn in allen Erhebungen stets neue Befragte mit überwiegend individuell zuzuordnenden Angaben auftreten wie beim ALLBUS. Zwar sind auch hier Speicherplatzeinsparungen möglich, sie werden aber zumindest teilweise wieder aufgewogen durch die Plattenplatzanforderungen, welche die DBMS-Strukturierung selbst beansprucht.[4]

Eine optimierte DBMS-Strukturierung in der gegenwärtig vorhandenen SIR-Version würde zudem eine inhaltlich ausgerichtete Aufbauhierarchie implizieren, z. B. Bundesländer Regierungsbezirke >Ortsgröße >Wahlverhalten etc. Dies hat zwei Konsequenzen: Erstens, Benut-

3 Es sollte an dieser Stelle angemerkt werden, daß die neuere Version 6 von SAS sehr viel weitergehende DBMS-Funktionen enthält, als sie ***Althainz*** 1987 zur Verfügung standen (vgl. ***Wallace*** 1989a, 1989b, *Kent* 1989a, 1989b). SAS ist in mancher Hinsicht gegenüber SPSS das mächtigere und flexiblere Programmpaket. SPSS ist demgegenüber in vieler Hinsicht einfacher anzuwenden bzw. benutzerfreundlicher (vgl. ***Hartmann*** 1988). Da zudem SAS voraussichtlich auch in der neuesten Version nicht auf allen Computeranlagen lauffähig sein wird (z.B. nicht auf den meisten Siemensgroßrechnern) und in den Sozialwissenschaften seltener verwendet wird als SPSS, kann SAS zur Zeit für die benutzerorientierte Haltung der ALLBUS-Daten nicht die erste Wahl sein. Es ist ferner an einigen Beispielen zu verfolgen, daß SPSS auf vielen Gebieten, wo SAS einen Vorsprung hatte, bald darauf ähnliche Programmerweiterungen einführte.

4 Zu dieser abzuspeichernden DBMS-Strukturierung gehört u.a. das "data dictionary" mit Metadaten-Tabellen über die Typen der in verschiedenen Einzeltabellen abgespeicherten Daten, ihre Schlüsselvariablen, Indices etc. Ferner müssen in allen Datentabellen Verknüpfungsmerkmale oder Schlüsselvariablen abgespeichert werden, mit denen die einzelnen Datenbankteile wieder zueinander in Beziehung gesetzt werden können (vgl. ***Uher*** 1989; ***Flory*** 1989).

zer, die mit dem komplexen Aufbau einer konkreten SIR-Datenbank nicht vertraut sind, können, abgesehen von der Komplexität des SIR-Retrievals als solchem, weniger effizient damit arbeiten. Zudem steigt die für weniger spezialisierte Benutzer ohnehin schwer überschaubare Fehleranfälligkeit einer Abfrage oder eines Retrievals. Zweitens, eine SIR-Datenbank sollte für ein **bestimmtes** Forschungsinteresse zugriffsmäßig optimiert werden. Das oben angegebene Aufbaubeispiel impliziert z. B., daß in der Regel die Benutzer einer solchen Datenbank an regionalisierten Daten unter dem Gesichtspunkt von Wahlanalysen interessiert sein dürften. Eine solche Annahme ist für den sehr unterschiedlichen Datenbestand und Benutzerkreis des ALLBUS generell nicht realistisch. Zwar ist es möglich, SIR-Datenbanken umzustrukturieren, so daß sie für andere Forschungsansätze optimiert sind, doch erfordert dies jeweils beträchtlichen Aufwand. Alle Benutzer mit Forschungsanliegen, für die eine SIR-Datenbank nicht vorstrukturiert ist, werden weniger effizient damit arbeiten, weil sie nicht entsprechend der bestehenden Aufbauhierachie auf die Daten zugreifen können. Bei anderen Datentypen, wie Unternehmens-, Panel- oder Netzwerkdaten, kann a priori rein formal eher eine allgemein vertretbare Vorstrukturierung eingeführt werden, weil zentrale Variablen existieren, von denen ausgehend sich Untertabellen und Standardzugriffspfade für die anderen Daten sinnvoll organisieren lassen (z. B. Fallnummern, für die mehrere Ereignisse vorliegen (Paneldaten) oder Fälle mit mehreren Beziehungstypen (Netzwerkdaten)).

Datensätze von der Struktur des ALLBUS sind für den standardmäßigen Zugriff und komplexere statistische Auswertungen relativ effizient in der herkömmlichen dateiorientierten Organisation zu halten, wenn die Daten unmittelbar "on line" in einer Form vorliegen können, welche für die vorwiegenden Auswertungsprogramme geeignet ist. Mehr oder weniger komplizierte Retrieval-Prozeduren, die einen Teildatensatz für komplexere statistische Analysen oder Weitergabe erst zugängüch machen, können so entfallen. Das bis auf weiteres standardmäßig für den allgemeinen Benutzerkreis vorzusehende Abspeicherungsformat sollte wie beim " **General Social Survey**" aus den USA am Grundmodell der rechteckigen Datenmatrix orientiert bleiben, was Zeitvergleiche und Zusammenfassen kleinerer Subgruppen für die Mehrzahl der Anwender erheblich erleichtert (vgl. auch *Davis/Smith* 1989: 1). Dies schließt die Entwicklung künftiger spezieller Datenbankangebote und interne Datenmanagement-Anwendungen nicht prinzipiell aus.

Das in den Sozialwissenschaften zur Zeit am weitesten verbreitete Programmpaket für statistische Analysen ist SPSS. Dieses System hat in den neueren Versionen von SPSSx einige neue Verarbeitungsmöglichkeiten integriert, welche die Datenhaltung erheblich erleichtern. Für die laufende Bearbeitung einer derartig großen Datei wie dem kumulierten ALLBUS bietet sich die Abspeicherung als **komprimierter** SPSSx-Systemfile an. Obgleich der komprimierte Systemfile neben den Daten auch alle Angaben des "data dictionary" enthält - etwa die Definition fehlender Werte sowie Variablen- und Wertelabel - beansprucht er weniger Plattenplatz

als der eingelesene Datensatz allein. Das Erstellen eines Systemfiles mit SPSS9 oder das Erstellen eines nicht komprimierten Systemfiles mit dem SPSSx erfordert dagegen sehr viel mehr Speicherplatz. Der kumulierte ALLBUS 1980-88 (S1795) belegt beispielsweise als komprimierter Systemfile 10 Mb gegenüber 55 Mb in nicht komprimierter Form. Einen Zeitgewinn bei Berechnungen unter Einsatz der nicht komprimierten Datei konnten wir auf unserem IBM-Rechner 4381 nicht feststellen. Zu diesem Vorteil verhältnismäßig effizienter Abspeicherung kommt bei Arbeiten mit Systemfiles im SPSSx eine merkliche Verkürzung der Bearbeitungs- und Rechenzeit gegenüber dem Arbeiten mit Setup und Rohdaten.

Durch einfache SELECT-, DROP- und KEEP-Angaben lassen sich darüberhinaus auf Systemfileebene weitere erhebliche Verkürzungen der Rechenzeit und eine Verringerung des für spezifische Analysen erforderlichen Speicherplatzes erzielen. *Terwey* 1988 gibt unter Verwendung von kumulierten ALLBUS-Daten einen Überblick, in welchem Umfang Reduzierungen des seitens der Anwender erforderlichen Plattenplatzes beim Erstellen eigener Analysedateien möglich sind. Hinsichtlich der Verkürzungen von erforderlichen Rechenzeiten kann an Beispielen festgestellt werden, daß ein vergleichbarer SPSS-Lauf je nach Menge der jeweils aktivierten Daten zwischen 355 oder 19 cpu-Sekunden benötigt (vgl. *Terwey* 1988). Ein Vergleich zwischen SPSSx und SIR beim Einsatz für die Paneldaten des britischen "General Household Survey" ergab, daß SPSSx nur zwischen 29 % und 63 % der Computerauslastung (accounüng units) erforderte wie SIR bei vergleichbaren Aufgabenstellungen *(Banks* 1987: 49).

Die ALLBUS-Daten liegen gegenwärtig im Zentralarchiv als Ausgangsbasis für Bearbeitung und Analysen zentral gespeichert vor und können für bestimmte Anwendungen selektiert und maßgerecht in einem neuen Systemfile von den Benutzern herausgeschrieben werden. Im Unterschied zum SIR-, INGRES- oder ORACLE-Retrieval sind diese Datenselektionen jedoch von jedem SPSS-Benutzer denkbar einfach durchzuführen. Die aufgrund der beschränkten statistischen Auswertungsmöglichkeiten vieler DBMS absehbare Notwendigkeit, einen für komplexere statistische Analysen geeigneten File aus der Datenbank erst zu extrahieren und in eine neue Programmumgebung zu exportieren, entfällt für alle im SPSSx enthaltenen oder programmierbaren Prozeduren.[5]

5 In der Regel werden die ALLBUS-Anwender für ihre statistischen Auswertungen im SPSS bequem einzusetzende fertige Prozeduren vorfinden. Eine Ausnahme sind in mancher Hinsicht die bisher im SPSSx enthaltenen Möglichkeiten, logistische Regressionen zu berechnen. Für dieses komplexe Verfahren sind beispielsweise im SAS mächtigere Prozeduren vorgegeben (vgl. *KühnellJagodzinskilTerwey* 1989, *Kühnel/Terwey* 1990). Im Release 4 des SPSSx sind jedoch mit dem dort eingeführten MATRIX-Programm auch schwierige statische Auswertungen neu programmierbar. Für die logistische Regression wurde von Steffen *Kühnel* mit MATRIX ein Programm erstellt, das auf sehr einfache Weise Daten einliest, binomiale und multinomiale logistische Regressionen berechnet und eine relativ einfach zu interpretierende Übersicht der relevanten statistischen Ergebnisse erstellt (vgl. den Beitrag von Steffen *Kühnel* in diesem Heft). Dies sei hier als ein Beispiel erwähnt, um zu zeigen, daß im SPSS heute Anwendungen realisierbar sind, die früher gar nicht in Erwägung gezogen werden konnten.

In Erwägung zu ziehen ist ferner, welche DBMS-Möglichkeiten das SPSSx bereits entwikkelt hat (z. B. INCLUDE, MATCH FILES, ADD FILES). Rohdateneingaben brauchen auch im SPSS nicht mehr notwendigerweise in einer rechteckigen Gesamtmatrix vorzuliegen, sondern können auf mehrere Einzeltabellen verteilt sein und/oder im Datentyp unterschiedlich strukturiert sein (siehe z. B. RECORD TYPE, FILE TYPE, FILE TYPE MIXED, FILE TYPE NESTED, INPUT PROGRAM). Damit sind im Bedarfsfall komplexere Abspeicherungsverfahren von Rohdaten realisierbar. Allerdings zeichnen sich beim Einsatz einiger dieser Möglichkeiten zum Einlesen sehr vielschichtig strukturierter Datensätze mit unterschiedlichen "record types" Grenzen einer einfachen und schnell zu programmierenden SPSS-Anwendung ab (siehe u.a. *Tagg* 1987, Projektgruppe 'Das sozioökonomische Panel' 1990).

Es ist zu erwarten, daß im SPSSx auch künftig versucht werden wird, weitere Datenmanagementfunktionen aufzunehmen, um die Marktposition des Systems zu stärken. Die jüngsten Programmerweiterungen in den neueren SPSSx-Versionen brachten beispielsweise zusätzliche Schnittstellen zu anderen Programmpaketen (BMDP, OSIRIS und SAS), neue Datenspezifikationsmöglichkeiten wie RENAME VARIABLES und die Möglichkeit, Dateien zu transponieren (FLIP).

2. Neuentwicklungen für ein verbessertes Daten- und Informationsmanagement

Bis auf weiteres ist es speichertechnisch vertretbar, die ALLBUS-Daten als einfache komprimierte Systemdateien im Rahmen des SPSSx zu halten, wodurch zugleich einem möglichst großen Benutzerkreis leichter Zugang gewährleistet werden kann. Das Einlesen in ein relationales DBMS könnte zur Zeit auch eine angemessene Verwaltung der Codebuchinformationen nicht gewährleisten. Die Datenbearbeitung würde dann zudem aufgrund der Verwendung von komplexeren Abfragestrukturen und DBMS-spezifischen Modifikationsanweisungen erheblichen Aufwand für die Einarbeitung von Bearbeitern und Anwendern bedeuten. Bei der von der ALLBUS-Arbeitsgruppe praktizierten SPSS-Bearbeitungsorganisation ist es dagegen noch vergleichsweise unkompliziert, von den verschiedenen Bearbeitern umfangreiche Datenmodifikationen temporär anzulegen und zu überprüfen. Nach eingehender Überprüfung werden diese Modifikationen zentral in der Datenbasis abgespeichert und können ggfs. auch auf einfache Weise wieder rückgängig gemacht werden, da in jedem Fall eine Kopie der Ausgangsdatenlage erhalten bleibt. "Rollback"-Korrekturen *(Schlageter/Stucky* 1983: 326-334), Systemfehler oder Zerstörungen der Datenbankintegrität sind hier kein die Datenbasis tangierendes Problem. Bereits bei der Datenaufbereitung im SPSS zeigt sich aber, daß eine große Kompetenz seitens der Bearbeiter erforderlich ist, die in der standardmäßigen SPSS-Ausbildung kaum vermittelt wird und mit den zur Verfügung stehenden finanziellen Mitteln nur

schwer auf dem Arbeitsmarkt zu rekrutieren ist. Unsere bisherigen Erfahrungen beim Einsatz von INGRES auf dem Großrechner zeigen zudem, daß die Inanspruchnahme des Rechners bei der Verwendung dieses relationalen Datenbanksystems ca. um den Faktor 5 steigen und somit die Rechengeschwindigkeit im Gesamtsystem von vornherein sinken würde.

Konkret ist jedoch angesichts der Komplexität der ALLBUS-Bearbeitung und der hierfür zur Verfügung stehenden Zeit der weitere Einsatz von technischen Hilfsmitteln zu wünschen. Neue Fazilitäten könnten insbesondere über Systeme laufen, die primär auf **Daten über die eigentlichen Daten** zugreifen (z.B. Variablennummern und andere Indices). Mit diesen **Metadaten** sind die eigentlichen Erhebungsdaten dann zu verwalten.

Ansatzpunkt für eine solche Neuentwicklung war zunächst eine Datei mit einer Strukturübersicht für die kumulierte Datenbasis. Diese Übersicht wurde zunächst als eine **Variablenkorrespondenzliste** realisiert, die im wesentlichen die Variablennummern aus der kumulierten Arbeitsdatei und den Einzelstudien enthielt. Diese Korrespondenzliste war als Übersicht zur Fragenkontinuität im ALLBUS anfangs nur für den Ausdruck bestimmt. Sie wurde aber so überarbeitet, daß sie als Metadatensatz von SPSSx eingelesen werden und nachfolgend als Basis für ein umfassenderes Datenmanagementsystem dienen kann (vgl. auch *Terwey* 1990). Die erste Anwendung dieses Systems lag zunächst noch in der Generierung eines neu strukturierten, druckfähigen Reports, einer am ALLBUS 1988 orientierten Variablenkorrespondenzliste, welche eine Übersicht über die im ALLBUS-Programm bereits vorher erhobenen, vergleichbaren Fragen enthielt. Obgleich dieser Report zwar im Layout für die Aufnahme in den Vergleichsdatenband zum ALLBUS 1988 (S1727) überarbeitet werden mußte, war dennoch ein beträchtlicher Zeitgewinn realisierbar.

Nachfolgend wurden weitere Metadatensätze mit Variablennummern und Indices angelegt, die über Schlüsselvariablen miteinander verknüpft werden können. Hierzu sind entsprechende SPSS-Steuerfiles anzulegen, welche die Relationen zwischen den Eingabedaten entweder temporär für weitere Arbeits schritte generieren oder einen permanenten neuen Datensatz erzeugen. Obgleich das Metadatensystem zur Zeit noch in der Entwicklung ist, konnten für die ALLBUS-Kumulationsstudie 1980-88 bereits weitere Anwendungen realisiert werden:

a) Mithilfe der Metadaten wurde die für die Kumulationsdatei relevante Teilmenge aus dem "file pool", der alle ALLBUS-Daten enthält, herausgeschrieben.

b) Diese Datenteilmenge wurde mittels Metadaten in eine neue, nach inhaltlichen Gesichtspunkten gegliederte Reihenfolge gebracht. Zu diesem Zweck wurden zahlreiche SPSS-Steueranweisungen durch SPSS weitgehend automatisch generiert.

c) Unter Einsatz des gleichen Systems wurden die für die Kumulation 1980-88 relevanten Codebuchtexte aus einem größeren Text-Pool extrahiert und entsprechend dem Datensatz neu geordnet.

d) Ferner wurden Programmteile weitgehend automatisch generiert, mit denen für bestimmte Variablen im Codebuch OSIRIS-Häufigkeiten und SPSS-Kreuztabellen berechnet werden konnten.

e) Neue Variablen-Labels mit maximal 40 Zeichen wurden aus den Metadaten selektiert und für die Kumulation 1980-88 in das Codebuch aufgenommen. Sie überschreiben dort die eher unbefriedigenden OSIRIS-Labels, die nur maximal 25 Zeichen lang sein können.

f) Die neuen Variablen-Labels stehen ferner nach der Übernahme in den ALLBUS-Filepool für die Generierung eines SPSS-Setups zur Verfügung.

In diesem Zusammenhang kommt ein anderer wichtiger Teil der Programminfrastruktur, die für den ALLBUS eingesetzt wird, zum Tragen. Wie bereits weiter oben erwähnt, ist im SPSSx keine dem SPSS9 vergleichbar bequeme Möglichkeit enthalten, aus Systemfiles ein Setup zu extrahieren. Der Hintergrund für den Wegfall dieser Möglichkeit ist vermutlich die Einführung von Exportfiles, die, aus SPSSx-Systemfiles erstellt, auch auf unterschiedlichen Computeranlagen gelesen werden können. Für den Archivdienst ist einer der Nachteile dieses Verfahrens, daß Anwender, die noch Analysen im SPSS9 durchführen, mit Exportfiles nicht arbeiten können. Selbst das kaum bekannte SPSSx-Programm UNPORT schafft in dieser Hinsicht keine brauchbare Abhilfe. Die SPSS-Setups für die neueren ALLBUS-Studien wurden deshalb, weitestgehend automatisiert, mithilfe von zwei speziellen Programmen erstellt, die aus dem OSIRIS DICT und dem DISPLAY DICTIONARY des SPSSx die relevanten Setupangaben in einer Form generieren, die auch unter SPSS9 lauffähig ist.

Durch die konsequente Verwendung von Metadaten konnte bei den laufenden Arbeiten eine Reihe von Fehlern, die beim manuellen Erstellen entsprechender Anweisungen erfahrungsgemäß auftreten, vermieden werden. Der Zeitaufwand sank insgesamt beträchtlich und viele manuelle Routinearbeiten entfielen. Da SPSS als Generator von Steueranweisungen für andere SPSS-Jobs ein in vieler Hinsicht einzusetzendes System ist, bleiben Möglichkeiten für andere ähnliche Erleichterungen bei der Erledigung von Routineaufgaben offen und nahezu beliebige Tabellenverknüpfungen für die Metadaten sind denkbar. Die erforderlichen Rechenzeiten für die Erstellung von Ausgaben für Daten- und Textmanagement sind sehr gering, ohne daß die Beanspruchung des Großrechners gegenüber anderen üblichen Bearbeitungen erhöht werden muß. Das System hat, obgleich es noch in der Entwicklung ist, bereits wichtige Arbeitshilfen gegeben. Es kann künftig prinzipiell, wenn entsprechende Arbeitsressourcen

zur Erstellung neuer Metadatentabellen verfügbar sind, auch für weitere Aufgaben wie Recherchen nach inhaltlichen Indices verwendet werden. Diese Indices können u.a. in der Benutzerberatung dazu dienen, für bestimmte Forschungsfragen relevante Variablen ausfindig zu machen (vgl. zu anderen bereits jetzt realisierbaren Recherchemöglichkeiten *Terwey* 1990).

Auch die Grundkonzeption der ALLBUS-Weitergabe an **externe** Benutzer sollte die Erhebungen als leicht zugängliche und benutzerfreundliche Datenbasis zur Verfügung stellen. Nur so können Benutzerzahlen erreicht werden, welche für eine allgemeine Bevölkerungsumfrage der Sozialwissenschaften wünschenswert sind. Erfahrungen in anderen Instituten zeigen, daß für ein häufiges datenbankmäßiges Retrieval, den bloßen Zugriff auf die Daten also, eigentlich ein spezieller Bearbeiter erforderlich wäre, der die Daten für die Forscher zugänglich macht. Aus anderen Archiven liegen ferner Erfahrungen mit großen Datensätzen vor, die sowohl in einer relationalen Datenbank als auch in einem einfachen "flat file" angeboten wurden. So ergab sich beim ICPSR (Interuniversity Consortium for Political and Social Research), daß die Nutzung beider Abspeicherungsformen zunächst annähernd gleichgewichtig war. Später zeigte sich jedoch, daß die Anwender der relationalen Datenbank zumeist die Daten als einfachen "flat file" nachbestellten. Der Aufwand für die seitens der Benutzer notwendigen Retrievals schränkte die Zeit, welche für die inhaltliche Auswertung der Daten erforderlich war, zu sehr ein.

Das für den ALLBUS eingeführte Daten- und Informationsmanagement ist sicherlich kein System, für das der Anspruch erhoben werden kann, prinzipiell besser zu sein als andere DBMS-Anwendungen. Wir haben, um es mit den Worten eines der Teilnehmer des 3. CESSDA-Expertenseminars auszudrücken, versucht, eine sehr komplizierte Aufgabe so einfach, schnell und benutzernah zu lösen, wie es noch möglich war. Sozialwissenschaftliche Einrichtungen mit anderen Daten und Aufgabenpriori täten haben im DBMS-Bereich für ihre komplexen Daten Anwendungen entwickelt, die sich für ihre Zwecke gut bewährt haben.

Literatur:

Abbenhardt, Helmut 1990:
Datenbanken leisten einen Beitrag zur Offenheit, in: Computerwoche 35: 31-35.

Althainz, Bettina 1987:
Vergleich des Datenbanksystems SIR mit dem Statistikprogrammpaket mit Datenmanagementfunktionen SAS anhand der Lebensverlaufsstudie des Sonderforschungsbereichs 3 der DFG, Diplomarbeit an der Technischen Fachhochschule Berlin.

Archilles, Albrecht 1989:
SQL. Standardisierte Datenbanksprache vom PC bis zum Mainframe, München und Wien: Oldenbourg.

Banks, Randy 1987:
SPSSx v. SIR. Some Considerations re. the GHS, in: The General Household Survey Newsletter No 4: 47-51.

Bender, Donald und Walter *Bien* 1988:
Kriterien für die Auswahl von Datenbanksystemen in den Sozialwissenschaften, in: ZA-Information 23: 113-118.

Chen, Peter Pin-Shan 1976:
The Entity-Relationship Model - Toward a Unified View of Data, in: ACM-Transactions on Database Systems, Vol. I: 9-36.

Chen, Peter Pin-Shan (ed.) 1980:
The Entity-Relationship Approach to System Analysis and Design, Amsterdam: North Holland.

Davis, James Allan und Tom W. *Smith* 1989:
General Social Surveys, 1972-1989, National Opinion Research Center (NORC ed.; Chicago)/The Roper Center for Public Opinion Research (Storrs, CT; distributor).

Eilers, Henk, Wim *Jansen* und Harry de *Volder* 1988:
SQL in der Praxis, Bonn et al.: Addison-Wesley.

Flory, Thomas S. 1989:
Commercial Relational Database Management Systems in a Social Research Environment: Expectations and Evaluation, in: ZA-Information 24: 128-138.

Gräf, Lorenz 1988:
Erfahrungen mit einer PC-Datenbank zur Verwaltung von Indikatoren in Panelerhebungen, in: ZA-Information 22: 31-44.

Hartmann, Petra 1988:
Wechsel der Analyseeinheit und File Handling - SPSS-X und SAS im Vergleich, in: Frank Faulbaum und Hans-Martin Uehlinger (Hg.), Fortschritte der Statistik-Software 1, Stuttgart: Gustav Fischer: 114-123.

Kent, Paul 1989a:
Using the New SQLProcedure in SAS Programs, in: SEUGI '89. Proceedings of the SAS European Users Group International Conference, Köln 09-12.05.1989:407-412.

Kent, Paul 1989b:
Views - Your Window on Data, in: SEUGI '89. Proceedings of the SAS European Users Group International Conference, Köln 09-12.05.1989:413-417.

Kühnel, Steffen M., Wolfgang *Jagodzinski* und Michael *Terwey* 1989:
Teilnehmen oder boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx, in: ZA-Information 25: 44-75.

Kühnel, Steffen M. und Michael *Terwey* 1990:
Einflüsse sozialer Konfliktlinien auf das Wahl verhalten im gegenwärtigen Vierparteiensystem der Bundesrepublik, in: Walter Müller, Peter Ph. Mohler, Barbara Erbslöh und Martina Wasmer (Hg.), Blickpunkt Gesellschaft: Einstellungen und Verhalten der Bundesbürger, Opladen: Westdeutscher Verlag: 63-94.

Mayr, Heinrich C, Klaus R. *Dittrich* und Peter C. *Lockemann* 1987:
Datenbankentwurf, in: Peter C. Lockemann und Joachim W. Schmidt (Hg.) Datenbank-Handbuch, New York, Berlin und Heidelberg: Springer: 486-557.

Mochmann, Ekkehard 1988:
DB2, INFORMIX, INGRES, NOMAD, ORACLE, REFLEX, SIR, SQL/DS oder was?, in: ZA-Information 23: 111-112.

Ponemereo, Ralph 1988:
The German Parliament Data Base and the Use of SIR, in: Jean-Philippe Genet (Hg.), Standardisation et dchange des bases des données historiques, Paris: Centre National de la Recherche Scientifique: 265-275.

Projektgruppe 'Das sozioökonomische Panel' 1990:
Das sozioökonomische Panel für die Bundesrepublik Deutschland nach fünf Wellen, in: Vierteljahreshefte zur Wirtschaftsforschung 1990/2: 141-151.

Schlageter, Gunter und Wolffried *Stucky* 1983:
Datenbanksysteme: Konzepte und Modelle, Stuttgart: Teubner.

Schlögell, Volker 1989:
Archive: Datenverwalter oder aktive Serviceinstitutionen? Anmerkungen zum RDBMS-Workshop im Zentralarchiv vom 28.-31.03.89, in: ZA-Information 24:121-127.

Tagg, S. K. 1987:
Taking SPSS-X to the Limit: Processing Raw GHS Data Files, in: ESRC-Data Archive Bulletin No. 38: S2-S4.

Terwey, Michael 1988:
The Data Base of the German "General Social Survey": ALLBUS, Papier präsentiert auf der "Cologne Computer Conference" am 07.09.-10.09.88 in Köln.

Terwey, Michael 1990:
The Data Base of the ALLBUS: Dealing with a Large, Cumulative Data Set, Papier präsentiert auf dem 3. CESSDA-Expertenseminar am 27.09.-30.09.90 in Köln.

Uher, Rolf 1989:
Der "Continuity Guide to the German Election Data Project 1953-1987" als relationale Datenbank realisiert, in: ZA-Information 24: 19-26.

Uher, Rolf 1990:
Continuity Guide to the German Election Data Project, 1953-1987. Realized as a Relational Database Concept, in: Conference Group on German Politics. Newsletter (March, 1990): 5-7.

Wallace, Emily 1989a:
Database Interfaces under the Version 6 Engine Architecture, in: SEUGI '89. Proceedings of the SAS European Users Group International Conference, Köln 09-12.05.1989:404-406.

Wallace, Jack 1989b:
Using New SAS Database Features and Options, in: SEUGI '89. Proceedings of the SAS European Users Group International Conference, Köln 09-12.05.1989: 392-403.

Wilke, Helmut und Wilfried ***Römer*** 1981:
SIR - Ein Programmpaket zur Verwaltung komplexer Datenstrukturen. Zeitbudget Daten jetzt als Datenbank erhältlich, in: ZA-Information 8: 16-19.

Zentralarchiv für empirische Sozialforschung (Hg.) 1990:
Daten der empirischen Sozialforschung. Datenbestandskatalog des Zentralarchivs mit Beschreibungen von Umfragen der empirischen Sozialforschung von 1945 bis 1990 und Daten der historischen Sozialforschung, Frankfurt und New York: Campus (im Druck).

Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen

von Wolfgang Ludwig-Mayerhofer

In Heft 25 bzw. 26 der ZA-Information haben ***Kühnel/Jagodzinski/Terwey*** (1989) bzw. ***Urban*** (1990) binäre und multinomiale Logit-Modelle zur Analyse kategorialer abhängiger Variablen mit zwei oder mehr Ausprägungen vorgestellt.

In dieser Arbeit soll gezeigt werden, wie in ähnlicher Weise auch multivariate Modelle geschätzt werden können, bei denen die abhängige Variable auf *Ordinalskalenniveau* gemessen wurde. Solche Variablen kommen in soziologischen Untersuchungen relativ häufig vor; manche Autoren vertreten die These, daß ein höheres Skalenniveau bei sozialwissenschaftlichen Messungen kaum erreichbar ist (z.B. ***Kriz*** 1973: 211; ***Heidenreich*** 1987: 359). Prüfverfahren für bivariate Zusammenhänge bzw. für Gruppenvergleiche mit ordinalskalierten Variablen sind zwar seit langem im sozialwissenschaftlichen Methodenarsenal bzw. in den entsprechenden Statistik-Paketen verfügbar. Adäquate *multivariate* Auswertungsverfahren, welche den simultanen Einfluß mehrerer unabhängiger Variablen zu prüfen erlauben, sind aber noch relativ unbekannt.

Beispielsweise wurde unlängst in einer kriminalsoziologischen Veröffentlichung ***(Geißler/Marißen*** 1988) untersucht, ob junge Frauen vor Strafgerichten (genauer gesagt: Jugendgerichten) gegenüber jungen Männern bevorzugt, also milder bestraft werden als diese. Die abhängige Variable, die gerichtliche Sanktion, wurde dem Schweregrad nach folgendermaßen geordnet: Verfahrenseinstellungen, sog. "Weisungen", "leichte Zuchtmittel" (Verwarnung, Geldbuße etc.), "schwere Zuchtmittel" (Arrest), Jugendstrafe mit Bewährung, Jugendstrafe ohne Bewährung, Anwendung von Erwachsenenstrafrecht. Die Autoren waren völlig zu Recht der Ansicht, daß man hierin zwar eine sinnvolle *Rangfolge* der Schwere der verhängten Sanktionen,[1] *nicht* aber eine Messung der Sanktionsschwere auf einer *Intervallskala* sehen kann, und haben daher Auswertungsverfahren für ordinalskalierte Variablen angewandt. Dabei haben sie sich jedoch auf die Berechnung einfacher und partieller Gamma-Koeffizienten beschränkt. Daß eine sinnvolle Analyse von zwei oder mehr unabhängigen Variablen auf diese Weise nicht möglich ist, dürfte offensichtlich sein.

---

[1] Die Einstufung von Sanktionen nach dem Erwachsenenstrafrecht als schwerste Reaktion, unabhängig von ihrer konkreten Ausformung, ist aus inhaltlichen Gründen allerdings problematisch, wie die Autoren selbst bemerken *(Geißler/Marißen* 1988, S. 516).

Die folgenden Ausführungen schließen an die Arbeiten von ***Kühnel et al.*** und von ***Urban*** an und stellen daher gleichfalls Logit-Modelle vor, jedoch in einer Form, welche die ordinale Information der abhängigen Variablen benutzt. Wie schon in den beiden vorherigen Arbeiten werden ausführliche Analysen eines einfachen Beispiels durchgeführt, um die Interpretation ordinaler Logit-Modelle möglichst gut nachvollziehbar zu machen (Abschnitt 1). Anschließend wird ein komplexeres Modell dargestellt (Abschnitt 2). Hier werden weitere Grundlagen, vor allem in inferenzstatistischer Hinsicht, diskutiert; da ordinale Logit-Modelle wie binäre und multinomiale Logit-Modelle auf Maximum-Likelihood-Schätzungen beruhen, ergeben sich viele Ähnlichkeiten mit den vorangegangenen Arbeiten. Es sollte allerdings nicht übersehen werden, daß auch andere Auswertungsmöglichkeiten für ordinalskalierte Variablen existieren; daher wollen wir diese im letzten Abschnitt (3) ganz kurz ansprechen. Der Anhang enthält Hinweise auf Programme zur Berechnung der hier vorgestellten Modelle, wobei auch die entsprechenden Angaben in den Arbeiten von ***Kühnel et al.*** und ***Urban*** aktualisiert werden.

1 Ein einfaches ordinales Logit-Modell: Staatsanwaltliche Reaktionen auf Jugendkriminalität

Unter einer ordinalskalierten Variablen soll hier - lose definiert - eine Variable mit zwei, in der Regel aber mehr Ausprägungen verstanden werden, welche in eine Rangfolge gebracht werden können, ohne daß genaue Aussagen über die Abstände zwischen diesen Ausprägungen möglich sind. Die Ausprägungen der Variablen entsprechen also einem "Mehr" bzw. "Weniger" des erfaßten Merkmals; dieses verschließt sich jedoch einer Quantifizierung im Sinne einer Intervallskala. Soll eine solche Variable als abhängige Variable in einem multivariaten Modell analysiert werden, so ließe sich zwar das von ***Urban*** dargestellte multinomiale Logit-Modell verwenden; dieses würde jedoch den ordinalen Charakter dieser Variablen nicht berücksichtigen. Außerdem würden in einem solchen Modell bei einer abhängigen Variablen mit r Ausprägungen für jede unabhängige Variable (r-1) ß-Koeffizienten geschätzt. Bei komplexen Modellen liegt dann sehr schnell eine kaum mehr sinnvoll interpretierbare Vielzahl von Koeffizienten vor. Auf der anderen Seite wurden und werden zwar sehr häufig lineare Regressionsmodelle auch bei ordinalskalierten Variablen angewandt; die Voraussetzungen dieses Modells - vor allem in inferenzstatistischer Hinsicht - sind jedoch bei solchen Variablen in der Regel verletzt.

Eine Alternative ist das hier vorgestellte ordinale Logit-Modell.[2] Dieses geht von den sog.

2 Theoretische Begründungen/Darstellungen dieses Modells finden sich vor allem in: *McKelvey/Zavoina* 1975; *McCullagh* 1980; *Maddala* 1983; *Agresti* 1983; ders. 1984; *Winship/Mare* 1984. Die hier gewählte Darstellung in Form kumulativer Logits ist vor allem von den Arbeiten von *Agresti* inspiriert.

"kumulativen Logits" aus, die sich am besten im Unterschied zu Logit-Modellen für kategoriale Variablen verdeutlichen lassen. In diesen, also im *binären oder multinomialen Logit-Modell,* werden die Wahrscheinlichkeiten der einzelnen Ausprägungen der abhängigen Variablen

$$p_i = P(y=i),\ i = 1, \ldots, r$$

jeweils *einzeln* zueinander in Beziehung gesetzt. Dabei genügt es, bei einer abhängigen Variablen mit r Kategorien das (logarithmisierte) Verhältnis von r-1 Kategorien zur verbleibenden (r-ten) Kategorie, mithin die Wirkungen der unabhängigen Variablen auf die Logits[3]

$$\ln \frac{p_i}{p_r},\ i = 1, \ldots, r-1$$

zu schätzen. Die Wahl der Bezugskategorie p_r ist dabei ebenso beliebig wie die Anordnung bzw. Indizierung der übrigen Kategorien, da es sich um eine nominalskalierte Variable handelt.

Dagegen verwendet das *ordinale Logit-Modell* als Zielvariable die r-1 Logits

$$(1a)\quad \ln \frac{p_1 + \ldots + p_i}{p_{i+1} + \ldots + p_r} = \ln \frac{p_1 + \ldots + p_i}{1 - (p_1 + \ldots + p_i)},$$

$i = 1, \ldots, r-1,$

oder anders ausgedrückt

$$(1b)\quad \ln \frac{P(y \le i)}{P(y > i)} = \ln \frac{P(y \le i)}{1 - P(y \le i)},$$

$i = 1, \ldots, r-1.$

Hier wird also die Wahrscheinlichkeit, in eine bestimmte oder eine "niedrigere" Kategorie (also $P(y \prec i)$), zur Wahrscheinlichkeit, in eine der "höheren" Kategorien zu fallen

---

[3] Ich verzichte im folgenden aus Gründen der Übersichtlichkeit auf die übliche Kennzeichnung der Schätzwerte durch ein "Hütchen". Sämtliche Formeln bzw. Berechnungen im folgenden Text beziehen sich auf Schätzwerte; wo nicht, ist dies jeweils explizit angegeben.

(also P(y > i)), in Beziehung gesetzt.[4] Dies ist aber nur sinnvoll, wenn "niedriger" bzw. "höher" eine inhaltliche Bedeutung haben, d.h., wenn die verschiedenen Kategorien eine Rangfolge darstellen. Insofern ist also vorausgesetzt, daß die abhängige Variable tatsächlich Ordinalskalenniveau hat und ihre Ausprägungen dieser Skala gemäß angeordnet sind. Jedoch gehen keine Annahmen hinsichtlich des "Abstandes" zwischen den Kategorien ein, wie sie einer Intervallskala entsprechen würden.

Ich möchte dies sogleich an einem einfachen Beispiel veranschaulichen, welches sich wie die eingangs erwähnte Untersuchung von *Geißler/Marißen* (1988) auf den Bereich der strafrechtlichen Sozialkontrolle Jugendlicher bzw. Heranwachsender[5] bezieht, allerdings auf die dem gerichtlichen Hauptverfahren vorgeschaltete Stufe der *staatsanwaltlichen Verfahrenserledigung.*[6] Dabei wollen wir die "traditionelle" Aufgabe der Staatsanwaltschaft, die Entscheidung über das Bestehen eines hinreichenden Tatverdachts, in den Hintergrund stellen, und untersuchen, wie die Staatsanwaltschaft verfährt, wenn sie entschieden hat, daß die vorliegenden Beweismittel ausreichen, den Tatverdächtigen als mutmaßlich Schuldigen zu betrachten.[7] Denn auch dann steht der Staatsanwaltschaft noch eine breite Palette von Entscheidungsalternativen offen. Die wichtigsten davon sind:

- *Einstellung des Verfahrens* nach § 45 Abs. 2 Nr. 2 JGG (Jugendgerichtsgesetz) *ohne weitere Reaktion.* Diese Verfahrenserledigung, die gewählt werden kann, wenn es sich

---

4 Diese wie spätere Formulierungen beziehen sich auf das Programm GAUSS (Prozedur "Ordered" im Modul "Quantal Response Models"), mit welchem die hier vorgestellten Analysen durchgeführt wurden. Gelegentlich werden statt dessen die "umgekehrten" Logits

$\ln ((p_{i+1} + \ldots p_t) / (p_1 + \ldots p_i))$, also

$\ln (P(y > i) / P(y \leq i))$

verwendet (z.B. bei *Agresti* 1983, 1984 und im BMDP-Modul "PR"). Diese Umkehrung drückt sich nur in der Vertauschung der Reihenfolge und der Vorzeichen für die nachfolgend dargestellten Regressionskonstanten $\alpha_1 \ldots \alpha_{t-1}$ und der Vertauschung der Vorzeichen für die ß-Koeffizienten aus (vgl. *McCullagh* 1980, S. 116).

5 Seit der Herabsetzung des Volljährigkeitsalters auf 18 Jahre gelten 18- bis unter 21jährige in strafrechtlicher Hinsicht als "Heranwachsende", bei denen Staatsanwalt und Richter zu prüfen haben, ob diese hinsichtlich ihrer "Reife" eher Jugendlichen oder Erwachsenen vergleichbar sind. In der Praxis hat sich ganz überwiegend durchgesetzt, daß Heranwachsende nach dem - zumeist milderen - Jugendstrafrecht behandelt werden.

6 Die Daten stammen aus einer Untersuchung nordrhein-westfälischer Jugendstaatsanwälte, welche im Teilprojekt C 1 des Sonderforschungsbereichs 227 "Prävention und Intervention im Kindes- und Jugendalter" an der Universität Bielefeld unter der Leitung von Peter-Alexis *Albrecht* durchgeführt wurde (vgl. dazu *Albrecht* 1990).

7 Die Einstellung des Verfahrens wegen nicht hinreichenden Tatverdachts folgt weitgehend anderen Regeln als die Entscheidung über die im folgenden untersuchten Erledigungsmöglichkeiten. Im übrigen ist sie bei Jugendlichen, zumal bei den im folgenden untersuchten Diebstahlsdelikten, vergleichsweise selten (vgl. dazu im einzelnen *Ludwig-Mayerhofer* 1990).

um ein geringfügiges Delikt handelt, hat also für den Beschuldigten (mit Ausnahme der Eintragung in das Erziehungsregister, welche u.U. bei erneuter Auffälligkeit von Nachteil sein kann) keinerlei Folgen.

- *Einstellung des Verfahrens* nach § 45 Abs. 2 Nr. 1 JGG, wenn eine *erzieherische Maßnahme* bereits "angeordnet" ist; solche Maßnahmen werden häufig von der Staatsanwaltschaft selbst - gegebenenfalls unter Einschaltung des Jugendamtes - durchgeführt, bestehen allerdings zumeist nur in einer schriftlichen oder mündlichen Ermahnung. Der Beschuldigte gilt, selbst wenn auf diese Weise eine offizielle "Reaktion" erfolgt ist, nicht als verurteilt.
- Antrag auf ein sog. *"vereinfachtes Verfahren"* nach § 76 JGG. Dieser Antrag kommt zwar der Anklageerhebung gleich, der Staatsanwalt muß jedoch bei der Hauptverhandlung nicht anwesend sein und ist es zumeist auch nicht; damit drückt er aus, daß sein Interesse an der gerichtlichen Sanktionierung eher gering ist.
- *Anklageerhebung vor dem Jugendrichter* als Einzelrichter. Der Einzelrichter kann höchstens Jugendstrafen bis zu einem Jahr verhängen, so daß diese Anklageform gewählt wird, wenn eine leichte oder mittlere Sanktionierung erwartet bzw. angestrebt wird.
- *Anklageerhebung vor dem Jugendschöffengericht* bzw. (in Ausnahmefällen) der Jugendstrafkammer. Diese Gerichte können Sanktionen bis zur Höchststrafe von 5 (bei Heranwachsenden auch bis zu 10) Jahren Jugendstrafe verhängen, so daß die Anklageerhebung vor diesen Gerichten erfolgt, wenn die Staatsanwaltschaft eine schwere Strafe durchsetzen will.

Diese fünf wichtigsten Erledigungsformen drücken also eindeutig eine - jedoch sicherlich nicht auf einer Intervallskala meßbare - Rangfolge der staatsanwaltlichen "Reaktionsschwere" aus: Vom völligen Reaktionsverzieht über die leichte Reaktion ohne gerichtliche Verurteilung bis zur Anklageerhebung, welche in den drei genannten Formen wiederum unterschiedlich schwere Sanktionsbestrebungen zum Ausdruck bringt.[8]

In einem multinomialen Logit-Modell würden (beliebige) vier der Kategorien dieser Variablen zur fünften Kategorie in Beziehung gesetzt und die Einflüsse unabhängiger

---

[8] Zwei weitere Möglichkeiten müssen erwähnt werden, welche in der staatsanwaltlichen Praxis - jedenfalls bei den hier untersuchten nordrhein-westfälischen Staatsanwälten - quantitativ eine marginale Rolle spielen: Der Staatsanwalt kann einmal das Verfahren auch nach Durchführung einer erzieherischen Maßnahme durch den Jugendrichter einstellen (§ 45 Abs. 1 JGG). Diese Fälle werden im folgenden zu den Einstellungen nach § 45 Abs. 2 Nr. 1 gerechnet. Ferner kann der Staatsanwalt bei Heranwachsenden einen Strafbefehl erlassen. Diese Erledigungsform haben wir zu den Anträgen auf das vereinfachte Verfahren hinzugerechnet. Eine eigenständige Analyse war wegen der geringen Häufigkeit dieser Erledigungen nicht möglich - womit auch eine Grenze der hier vorgestellten Modelle angesprochen ist.

Variablen auf die entsprechenden vier Logits (also Logarithmen der Proportionen) geschätzt. Die Tatsache, daß diese fünf Kategorien eine Rangordnung darstellen, wäre für dieses Modell völlig irrelevant. Das ordinale Logit-Modell geht dagegen von folgender Überlegung aus: Wenn die Kategorien der abhängigen Variablen - in diesem Fall fünf - eine Rangordnung bilden, dann sollte eine unabhängige Variable das Verhältnis
(1. Kategorie) / (2. + 3. + 4. + 5. Kategorie)
in ähnlicher Weise beeinflussen wie das Verhältnis
(1..+ 2. Kategorie) / (3. + 4. + 5. Kategorie) bzw.
(1. + 2. + 3. Kategorie) / (4. + 5. Kategorie) bzw.
(1. + 2. + 3. + 4. Kategorie) / (5. Kategorie).
Diese Zusammenfassung jeweils nebeneinanderliegender Kategorien impliziert aber, daß diese Kategorien auch im Sinne einer Rangordnung "zusammengehören". Der Ausdruck "kumulative Logits" bezieht sich auf den Sachverhalt, daß hier jeweils die "zusammengehörigen" Wahrscheinlichkeiten kumuliert werden.

Die abhängige Variable hat also ebenso viele Ausprägungen wie in einem multinomialen Logit-Modell, so daß wie in jenem Modell (r-1) Regressionskonstanten $\alpha_1 \ldots \alpha_{r-1}$ geschätzt werden. Weil das ordinale Logit-Modell aber unterstellt, daß der Einfluß der unabhängigen Variablen auf die kumulativen Logits auf jeder Stufe gleich ist, wird für jede unabhängige Variable nur ein einziges Regressionsgewicht geschätzt. Das Modell lautet also bei k unabhängigen Variablen

$$(2)\quad \ln \frac{p_1 + \ldots + p_i}{1 - (p_1 + \ldots + p_i)} = \alpha_i - \beta_1 * x_1 - \ldots - \beta_k * x_k,$$

für i = 1 ... r-1,

so daß also insgesamt (r-1) Regressionsgleichungen geschätzt werden, die sich jedoch nur hinsichtlich der Konstanten unterscheiden.

Im folgenden soll zunächst als Beispiel ein ordinales Logit-Modell mit nur einer unabhängigen Variablen dargestellt werden. Da die staatsanwaltliche Reaktion unterschiedlich starke Sanktionsbestrebungen zum Ausdruck bringt, können wir vermuten, daß sie mit der *Schwere des Tatvorwurfs* zusammenhängt, für die wir als Indikator hier zunächst die *Anzahl der Delikte* verwenden, welche den Beschuldigten vorgeworfen werden. Wir erwarten also, daß die staatsanwalüiche Reaktion umso schwerer ausfällt, je mehr Delikte der Beschuldigte mutmaßlich verübt hat. Wollte man nur diesen bivariaten Zusammenhang analysieren, wäre natürlich die Berechnung eines ordinalen Zusammenhangsmaßes ausrei-

chend. Wie im Anschluß gezeigt wird, läßt sich das ordinale Logit-Modell aber auf (jedenfalls grundsätzlich) beliebig viele unabhängige Variablen erweitern; das bivariate Beispiel wird hier gewählt, um zunächst einen Einstieg zu ermöglichen.

Tabelle 1: Zusammenhang zwischen der Anzahl der vorgeworfenen Delikte und der Verfahrenserledigung durch den Jugendstaatsanwalt (Nordrhein-Westfalen 1987, nur Diebstahlsdelikte)

	ZAHL DER DELIKTE				
ERLEDIGUNG	1	2	3-4	5 und mehr	N
Einstellung ohne weitere Reaktion	1190 25.3% *24.8%*	12 4.1% *10.9%*	3 1.8% *4.4%*	1 .6% *1.7%*	1206
Einstellung mit Reaktion	626 13.3% *13.4*	18 6.1% *7.8%*	9 5.4% *3.5%*	3 1.9% *1.4%*	656
Vereinfachtes Verfahren	500 10.6% *11.0%*	24 8.1% *7.8%*	21 12.5% *3.9%*	3 1.9% *1.7%*	548
Anklage vor dem Jugendrichter	1993 42.4% *42.4%*	166 56.1% *53.6%*	75 44.6% *48.2%*	53 33.5% *31.0%*	2287
Anklage vor dem Jugendschöffenger.	390 8.3% *8.4%*	76 25.7% *19.9%*	60 35.7% *40.0%*	98 62.0% *64.3%*	624
N	4699	296	168	158	5321

1. Prozentangabe (Normalschrift): Beobachtete Anteilswerte
2. Prozentangabe (Kursivschrift): Durch ordinales Logit-Modell geschätzte Anteilswerte

In *Tabelle 1* ist der Zusammenhang zwischen staatsanwaltlicher Verfahrenserledigung und Zahl der (nach Ansicht der Staatsanwaltschaft!) verübten Delikte in Form einer Kreuztabelle dargestellt. Wir können erkennen, daß die Ausgangsvermutung in der Tat zutrifft: Je größer die Anzahl der Delikte, desto seltener werden die Verfahrenseinstellungen und desto häufiger werden die Anklagen. Ein ordinales Logit-Modell für den in Tabelle 1 dargestellten Zusammenhang (wobei die vier Ausprägungen der unabhängigen Variablen

als 0, 1,2 und 3 kodiert wurden[9]) schätzt also 4 α-Koeffizienten, weil die abhängige Variable 5 Ausprägungen aufweist, sowie einen ß-Koeffizienten für den Einfluß der unabhängigen Variablen. Wir erhalten folgende Schätzung für die Koeffizienten:

α_1: -1.1083; α_2: -0.4808; α_3: -0.0317; α_4: 2.3854
ß: 0.9910

Im folgenden soll nun dargestellt werden, wie sich diese Ergebnisse in sinnvolle Aussagen über den Zusammenhang zwischen der Anzahl der Delikte und der staatsanwaltlichen Entscheidung übersetzen lassen.

Die *Konstanten* α_1 *bis* α_4 lassen sich wie in jedem Regressionmodell auffassen als Schätzung des Wertes der abhängigen Variablen für diejenigen Fälle, die in der (bzw. den) unabhängigen Variablen den Wert 0 aufweisen. Die α-Koffizienten enthalten also folgende Modell-Aussage über die kumulativen Logits: Für einen Fall, welcher in der *unabhängigen Variablen den Wert 0 aufweist* (wenn also der Beschuldigte 1 Delikt begangen hat), beträgt die Schätzung für In $(p_1 / (1-p_1))$ *-1.11* bzw. für $(p_1 / (1-p_1))$ $\exp(-1.11) = 0,33$. Das Modell formuliert also, daß das Verhältnis der Wahrscheinlichkeit, in die erste Kategorie der abhängigen Variablen (Verfahrenseinstellung ohne weitere Reaktion), zur Wahrscheinlichkeit, in die übrigen Kategorien zu fallen, eindeutig negativ ist, genauer gesagt, etwa 1:3 beträgt (was sich auch für die empirischen Daten leicht nachrechnen läßt). $\mathrm{Ln}\,((p_1 + p_2) / (1 - (p_1 + p_2)))$ ist mit *-0,48* zwar größer, dies entspricht aber immer noch einem negativen Verhältnis von exp(-0,48) = 0,62, also ca. 1:1,5. Erst das Verhältnis $(p_1 + p_2 + p_3) / (1 - (p_1 + p_2 + p_3))$ ist praktisch ausgeglichen, und die Relation $(p_1 + p_2 + p_3 + p_4) / (1 - (p_1 + p_2 + p_3 + p_4))$ ist positiv, und zwar mit exp(2,39) = 10,9 sehr erheblich. Dies läßt sich auch an den empirisch beobachteten Daten nachvollziehen: Da die 5. Kategorie nur sehr gering besetzt ist, ist die Chance, nicht in die 5. (sondern in Kategorie 1 bis 4) zu fallen, sehr groß. Dieser Wert wird also insbesondere durch die große Besetzung von Kategorie 4 bewirkt.

Der *ß-Koeffizient von 0,99* besagt nun laut Gleichung (2), daß sich die vier kumulativen Logits mit jeder Zunahme der unabhängigen Variablen um eine Einheit um den Betrag von -0,99 ändern. D.h., das Modell schätzt, daß für alle 4 kumulativen Proportionen der abhängigen Variablen (also Zusammenfassungen von (nebeneinanderliegenden) "weniger

---

9 Die Zusammenfassung von 3 und 4 sowie von 5 und mehr Delikten zu je einer Kategorie erfolgte nicht nur aus Gründen der Überschaubarkeit des Beispiels. Nach unserer Kenntnis des Datenmaterials können wir davon ausgehen, daß die Staatsanwälte kaum zwischen der Verübung von 3 oder 4 und ebensowenig zwischen 5 oder mehr Delikten unterscheiden. Die verwendete manifeste Skala dürfte also der zugrundeliegenden latenten Schwereeinschätzung durch die Staatsanwälte entsprechen.

eingriffsintensiven" und "schwereren" staatsanwaltlichen Reaktionen) das *Verhältnis* ersterer zu letzteren abnimmt (oder umgekehrt: das Verhältnis letzterer zu ersteren zunimmt), je mehr Delikte verübt worden sind; die "milderen" Reaktionen werden also seltener, die "schwereren" Reaktionen häufiger. Dies deckt sich mit unserer Ausgangsvermutung; gleichwohl ist diese Aussage des Modells recht abstrakt, und wir werden gleich darauf eingehen, wie man sie besser handhabbar machen kann. Doch zuerst sollen auf dieser Ebene die Modellschätzungen mit den empirischen Ergebnissen verglichen werden.

Tabelle 2: Empirische (obere Zeile) und geschätzte (untere Zeile) kumulative Logits für den Zusammenhang zwischen Anzahl der Delikte und staatsanwaltlicher Reaktion (vgl. Tab. 1)

	ZAHL DER DELIKTE			
KUMULATIVE LOGITS	1	2	3 bis 4	5 und mehr
$\ln \frac{p_1}{1-p_1}$	-1,08	-3,16	-4,01	-5,06
	-1,11	-2,10	-3,09	-4,08
$\ln \frac{p_1+p_2}{1-(p_1+p_2)}$	-0,46	-2,18	-2,56	-3,65
	-0,48	-1,47	-2,46	-3,45
$\ln \frac{p_1+p_2+p_3}{1-(p_1+p_2+p_3)}$	-0,03	-1,50	-1,41	-3,07
	-0,03	-1,02	-2,01	-3,00
$\ln \frac{p_1+p_2+p_3+p_4}{1-(p_1+p_2+p_3+p_4)}$	2,40	1,06	0,59	-0,49
	2,39	1,39	0,40	-0,59

In *Tabelle 2* sind die empirischen und die durch das Modell geschätzten kumulativen Logits in Abhängigkeit von den verschiedenen Ausprägungen der unabhängigen Variablen

"Zahl der Delikte" eingetragen. Um noch einmal zu verdeutlichen: Die Werte für die *geschätzten* kumulativen Logits sind in der ersten Spalte die Schätzer für α_1 bis α_4; in der zweiten bis vierten Spalte wird in jeder Zeile von dem Wert in der vorherigen Spalte der Betrag 0,99 abgezogen. Die *empirischen* kumulativen Logits lassen sich direkt aus den Häufigkeiten berechnen; so ergibt sich der Wert von -1,08 in der ersten Zeile der Tabelle aus ln (1190 / (4699 - 1190)), der Wert von -0,46 in der Zeile darunter aus ln ((1190 + 626) / (4699 - (1190 + 626))), usw. Insgesamt wird man feststellen können, daß das Logit-Modell eine brauchbare Modellierung der beobachteten Werte darstellt, wenngleich der Modellfit am "Rand", also in der ersten und vierten Spalte, besser ist als in den übrigen beiden Spalten.

Anstelle der kumulativen Logits ließen sich natürlich auch die nicht-logarithmisierten kumulativen Proportionen angeben; allerdings ist auch diese Betrachtungsweise nicht sehr anschaulich, und auch die Interpretation des ß-Koeffizienten ist immer noch sehr abstrakt. Sie würde lauten: Die vier geschätzten kumulativen Proportionen betragen innerhalb jeder Zeile, also mit jeder Zunahme der unabhängigen Variablen um eine Einheit, jeweils das exp(-0,99) = 0,37-fache der "vorherigen" geschätzten Proportion. Das heißt, mit jeder Zunahme der unabhängigen Variablen beträgt das *Verhältnis* der "milderen" zu den "strengeren" staatsanwaltlichen Reaktionen nur mehr ein gutes Drittel im Vergleich zur vorangegangenen Ausprägung der unabhängigen Variablen. Beispielsweise - um uns auf die ersten beiden Spalten der ersten Zeile in Tabelle 2 zu beziehen (geschätzte Werte!) - ist exp(-1,11) = 0,33, exp(-2,10) = 0,123; das Verhältnis von 0,123 zu 0,33 beträgt aber genau 0,37.[10] Wir können also unsere Aussage, daß die milderen Reaktionen umso seltener werden, je mehr Delikte der Beschuldigte begangen hat, auf diese Art und Weise quantifizieren; aber die auf Proportionen oder Verhältnisse bezogene Formulierung ist sicherlich nicht sehr eingängig.

Eine wesentlich besser verständliche Möglichkeit, die Aussagen des Modells darzustellen, ergibt sich aus einer Transformation, aus welcher die Schätzungen des Modells für die (bedingten) Wahrscheinlichkeiten innerhalb jeder Spalte von Tabelle 1 resultieren. Wir können jede Wahrscheinlichkeit $P(y = i)$ als Differenz zwischen $P(y \leq i)$ und $P(y \leq i\text{-}1)$ auffassen; beispielsweise ist die Wahrscheinlichkeit, in die 4. Kategorie zu fallen, gleich der Wahrscheinlichkeit in die 1. bis 4. Kategorie zu fallen, abzüglich der Wahrscheinlichkeit, in die 1. bis 3. Kategorie zu fallen.

---

[10] Ebenso ließe sich natürlich mit dem Kehrwert operieren: Das Verhältnis der "höheren" zu den "niedrigeren" kumulativen Proportionen beträgt jeweils das exp(0,99) = 2,7fache.

Aus

$$(3)\ \Pr(Y \leq i \mid x) = \frac{\exp(\alpha_i - \beta x)}{1 + \exp(\alpha_i - \beta x)}$$

ergibt sich mithin (***Anderson/Philips*** 1981, S. 23, ***Winship/Mare*** 1984, S. 515):

$$(4)\ \Pr(Y = i \mid x) = \frac{\exp(\alpha_i - \beta x)}{1 + \exp(\alpha_i - \beta x)} - \frac{\exp(\alpha_{i-1} - \beta x)}{1 + \exp(\alpha_{i-1} - \beta x)}$$

Dabei ist $\alpha_0 = -\infty$ und $\alpha_r = +\infty$, so daß im Fall der ersten Kategorie der abhängigen Variablen der Teil von Formel 4 *nach* dem Minuszeichen lautet:

$\exp(\alpha_{i-1} - \beta x) / (1 + \exp(\alpha_{i-1} - \beta x)) =$
$\exp(\alpha_0 - \beta x) / (1 + \exp(\alpha_0 - \beta x)) = 0,$

während sich im Fall der letzten (r-ten) Kategorie für den Teil *vor* dem Minuszeichen ergibt:

$\exp(\alpha_i - \beta x) / (1 + \exp(\alpha_i - \beta x)) =$
$\exp(\alpha_r - \beta x) / (1 + \exp(\alpha_r - \beta x)) = 1.$

Mithin errechnet sich z.B. für die erste Zeile der Kreuztabelle

$\Pr(y=1 \mid x=0) = (\exp(-1{,}11) / (1 + \exp(-1{,}11))) - 0 = 0{,}248,$

was mit 100 multipliziert als geschätzter Prozentanteil in *Tabelle 1 als unterer Wert* für diese Zeile eingetragen ist. Für die Zeile (3,3) (also Wahrscheinlichkeit eines vereinfachten Verfahrens bei drei oder vier Delikten - was numerisch mit 2 codiert wurde!) ergibt sich

$(\exp(-0{,}03 - 0{,}99*2) / (1 + \exp(-0{,}03 - 0{,}99*2))) -$
$(\exp(-0{,}48 - 0{,}99*2) / (1 + \exp(-0{,}48 - 0{,}99*2))) = 0{,}039,$

usw. In ähnlicher Weise kann auch bei mehreren unabhängigen Variablen für jede Konstellation von Werten die Wahrscheinlichkeit der verschiedenen "Outcomes" berechnet

werden.[11] Im gegenwärtigen einfachen Beispiel ist jedenfalls auch auf dieser Ebene die weitgehend recht gute Übereinstimmung von Modell und beobachteten Daten zu erkennen.

2 Ein komplexeres Modell: Inferenzstatistische Aspekte und weitere Interpretationen

Aufgrund seiner im Vergleich zu einem multinomialen Logit-Modell einfachen Struktur erlaubt das ordinale Logit-Modell verhältnismäßig komplexe Analysen, ohne daß die Ergebnisse unüberschaubar werden. Um bei unserem Beispiel zu bleiben: Die staatsanwaltliche Entscheidung hängt keineswegs alleine von der Zahl der Delikte ab. Hinzu kommen weitere Indikatoren für die "Schwere des Delikts", allen voran die Höhe des durch das bzw. die Delikt(e) bewirkten Schadens, und außerdem Faktoren, die sich auf den Tatverdächtigen beziehen, vor allem seine bisherige strafrechtliche Auffälligkeit, möglicherweise aber auch Merkmale wie Geschlecht, Alter, Nationalität, Arbeitslosigkeit etc. Es liegt also eine genuin "multivariate" Entscheidungssituation vor: Der Staatsanwalt berücksichtigt unter Umständen eine ganze Reihe von Faktoren, um die seiner Meinung nach angemessene Form der Verfahrenserledigung wählen zu können.

Tabelle 3 enthält die Ergebnisse eines multivariaten Modells mit einer erheblichen Zahl von Prädiktoren. Von den 17 geprüften Variablen (darunter zwei, welche bei drei Ausprägungen in je zwei Dummy-Variablen zerlegt wurden, so daß insgesamt 19 Prädiktoren getestet wurden) sind angesichts der großen Stichprobe nur drei nicht signifikant. Während ein multinomiales Logit-Modell bei der gleichen Anzahl von signifikanten Prädiktoren 68 Koeffizienten (einschließlich Konstanten) schätzen würde, enthält das ordinale Logit-Modell nur 20 Koeffizienten, ist also wesentlich übersichtlicher. Um die Koeffizienten untereinander vergleichbar zu machen, wurden sie auch mit den Standardabweichungen der jeweiligen erklärenden Variablen multipliziert (rechte Spalte in Tabelle 3). Sie sind gruppiert nach Merkmalen, die sich auf das Delikt, und solchen, die sich auf den Tatverdächtigen beziehen; innerhalb dieser Gruppen sind sie nach der Größe der standardisierten Koeffizienten geordnet. Hinzu kommt der Strafantrag des Opfers als "verfahrensbezogenes" Merkmal. Zum Vergleich sind auch die Ergebnisse eines linearen Regressionsmodells angegeben, welches mit Ausnahme der Kategorie "Sonstiger Diebstahl" die gleichen Einflüsse als signifikant (und auch die gleichen als nicht signifikant) ausweist wie das ordinale Logit-Modell.

---

[11] In diesem Fall tritt natürlich an die Stelle von ßx (also der einzigen unabhängigen Variablen, gewichtet mit dem dazugehörigen Regressionskoeffizienten) $\Sigma ß_p x_p$ (p = 1 ...k), also die *Summe* der mit den ß-Koeffizienten gewichteten Ausprägungen der unabhängigen Variablen.

Tabelle 3: Multivariates ordinales Logit-Modell für die staatsanwaltliche Entscheidung in Jugendstrafverfahren (in Klammern zum Vergleich: Entsprechende Werte eines linearen Regressionsmodells)

Variable	Koeff.	S. E.	t	Stand. Koeff.
α_1	-0,720	0,1064	-6,77	
α_2	0,098	0,1066	0,92	
α_3	0,780	0,1085	7,19	
α_4	4,723	0,1393	33,91	
(Lineare Regressionskonstante	2,502	0,0591)		
Tatmerkmale:				
Schadenshöhe	0,357	0,0182	19,59	0,788
	(0,171	0,0090	19,01	0,272)
Anzahl der Delikte	0,528	0,0557	9,48	0,336
	(0,135	0,0258	5,24	0,062)
Einbruch +	0,819	0,1013	8,08	0,305
	(0,225	0,0521	4,32	0,060)
Sonstiger Diebstahl +	-0,294	0,0792	-3,71	-0,119
	(-0,079	0,0430	-1,84	-0,023)
Haupttäter	0,466	0,0635	7,34	0,223
	(0,268	0,0337	7,94	0,092)
Tatverdächtigenmerkmale:				
Vorbelastung	0,483	0,0213	22,70	0,779
	(0,201	0,0102	19,76	0,235)
Reue	-0,659	0,0674	-9,78	-0,330
	(-0,373	0,0360	-10,37	-0,135)
Alter: Heranwachsender ++	0,570	0,0618	9,23	0,271
	(0,265	0,0338	7,86	0,091)
Arbeitslosigkeit	0,761	0,0933	8,15	0,243
	(0,305	0,0492	6,20	0,070)
Volles Geständnis +++	-0,378	0,0769	-4,91	-0,168
	(-0,193	0,0418	-4,63	-0,062)
Keine Beschuldigtenvernehmung +++	-0,450	0,1168	-3,85	-0,113
	(-0,186	0,0652	-2,85	-0,034)
Geschlecht: Weiblich	-0,302	0,0660	-4,57	-0,133
	(-0,173	0,0354	-4,90	-0,055)
Auffälligkeit beim ersten Zugriff	0,578	0,1287	4,49	0,123
	(0,259	0,0694	3,73	0,040)
Nationalität: Türkisch	0,361	0,0978	3,69	0,102
	(0,212	0,0522	4,06	0,043)
Abweichender Wohn-/Aufenthaltsort	0,335	0,1203	2,78	0,078
	(0,219	0,0636	3,45	0,037)
Strafantrag	0,337	0,0747	4,52	0,128
	(0,165	0,0406	4,06	0,045)

Nicht signifikant: Sonderschulbesuch; Schadenswiedergutmachung;
Vorhandensein/Fehlen von Tatzeugen.

Devianz:	11874,21
Devianz des Nullmodells:	15355,15
Pseudo-R^2:	0,23
(R^2 im linearen Modell	0,42)

+: Referenzkategorie: Ladendiebstahl
++: Heranwachsende: 18- bis 20jährige im Vergleich zu unter 18jährigen
+++: Referenzkategorie: Bestreiten des Tatvorwurfs (völlig oder teilweise)

Die Ergebnisse entsprechen allesamt den Erwartungen:[12]

Tatmerkmale: Je höher der durch das bzw. die Delikt(e) bewirkte Schaden und je größer die Anzahl der Delikte, desto schwerer die staatsanwaltliche Reaktion. Hat ein Beschuldigter einen Einbruch verübt, so führt das im Vergleich zur Referenzkategorie Ladendiebstahl zu einer schwereren, ein "sonstiger" Diebstahl dagegen zu einer leichteren Reaktion. Ebenso werden Beschuldigte, welche ihr(e) Delikt(e) als Haupttäter in einer Gruppe verübt haben, strenger sanktioniert als solche, welche die Tat(en) alleine oder nur als "Mitläufer" begangen haben.

Tätermerkmale: Den deutlichsten Einfluß hat mit großem Abstand die "strafrechtliche Vorbelastung", also die Anzahl der bereits früher gegen den Beschuldigten durchgeführten Strafverfahren. Aber auch mehrere andere Merkmale sind von Bedeutung. Staatsanwälte bevorzugen einmal Beschuldigte, die - beispielsweise bei der polizeilichen Vernehmung - Reue gezeigt haben; ebenso wirkt sich ein volles Geständnis "mildernd" auf die staatsanwaltliche Reaktion aus.[13] Hat umgekehrt ein Beschuldigter beim "ersten Zugriff", also

---

12 Begründungen der Hypothesen sowie ausführlichere Darstellungen der verwendeten Variablen finden sich bei *Ludwig-Mayerhofer* (1990). Dort werden die hier analysierten Daten, aber auch Daten zu weiteren Delikten, vor allem hinsichtlich der Dichotomie "Einstellung" vs. "Anklage" untersucht. Abweichungen der Fallzahlen der zitierten von denen der vorliegenden Arbeit ergeben sich dadurch, daß für die gegenwärtige Arbeit aus Gründen der Vereinfachung Fälle mit fehlenden Werten ausgeschlossen wurden.

13 Die mildere Reaktion bei fehlender Vernehmung des Beschuldigten dürfte sich so deuten lassen, daß die Staatsanwälte dann, wenn sie aus anderen Gründen eine milde Reaktion wählen, auf eine Beschuldigtenvernehmung verzichten.

z.B. bei der Festnahme durch die Polizei, sich auffällig verhalten, z.B. Widerstand geleistet oder einen Fluchtversuch unternommen, ist dies für die Staatsanwälte Anlaß für eine härtere Reaktion. Wir finden aber auch eine soziale Selektivität: Neben dem Alter führen Arbeitslosigkeit, ein "auffälliger" Wohn-/Aufenthaltsort (Erziehungsheim, ohne festen Wohnsitz) sowie türkische Nationalität des Beschuldigten zu einer schwereren Reaktion. Auch eine "mildere" Behandlung weiblicher Beschuldigter im Vergleich zu männlichen Beschuldigten läßt sich feststellen, wenngleich der Effekt des Geschlechts angesichts des sehr erheblichen Einflusses anderer Variablen wie der Schadenshöhe oder der strafrechtlichen Vorbelastung, aber auch z.B. der Arbeitslosigkeit, eher nachrangig ist.

Der *Strafantrag* des Opfers soll eigentlich nur die Funktion haben, eine Verfahrenseinstellung wegen "fehlenden öffentlichen Interesses" zu verhindern. Offenbar sind die Staatsanwälte aber auch darüber hinausgehend geneigt, den Strafverfolgungsinteressen der Opfer entgegenzukommen, indem sie zu einer "strengeren" Reaktion greifen, wenn der Strafantrag vorliegt.

Da auch ordinale Logit-Modelle nach dem Maximum-Likelihood-Verfahren geschätzt werden, lassen sich inferenzstatistische Aussagen sowie solche über die Gesamterklärungskraft des Modells völlig analog zu den Arbeiten von *Kühnel et al.* bzw. *Urban* formulieren:[14]

Signifikanz einzelner Parameter: Die Standardabweichungen der Koeffizienten lassen sich aus der Informationsmatrix berechnen; sie werden von den verwendeten Programmen ebenso wie die t-Werte (Verhältnis der Koeffizienten zu ihrer Standardabweichung) routinemäßig ausgegeben. Im Beispiel sehen wir, daß mit einer Ausnahme ("abweichender Wohn-/Aufenthaltsort") die Koeffizienten im herkömmlichen Sprachgebrauch sogar als "höchst signifikant" zu gelten haben, da die Werte in der Spalte "t" über 3,291 hegen.[15]

Signifikanz und Erklärungskraft des Gesamtmodells: Wie bei binären und multinomialen Logit-Modellen folgt die Differenz zwischen der Devianz des Nullmodells, also des Modells, welches nur die Regressionskonstanten schätzt, und der Devianz des vollen Modells, also des Modells mit allen signifikanten Prädiktoren, einer Chi-Quadrat-Verteilung mit k Freiheitsgraden (k = Anzahl der geschätzten Koeffizienten einschließlich der α-Koeffizienten). Im vorliegenden Fall ergibt sich ein Chi-Quadrat von 3473,02, welches bei 20

---

14 Vgl. hierzu im Kontext ordinaler Logitmodelle vor allem *McKelvey/Zavoina* 1975.

15 Vgl. aber die Hinweise bei ***Guttman*** 1977, S. 91 f., auf die Sinnlosigkeit solcher Bezeichnungen. Bei der Prüfung der Variablen wurde hier wie üblich ein Signifikanzniveau von 0,05 zugrunde gelegt.

Freiheitsgraden das gewählte Modell als signifikant erklärungskräftiger als das Null-Modell ausweist (was angesichts der großen Fallzahl nicht überraschend ist). Ebenso läßt sich als Maß für die Gesamterklärungskraft des Modells Pseudo-R^2 als
1 - (Devianz volles Modell / Devianz Nullmodell) berechnen. Im vorliegenden Fall beträgt Pseudo-R^2 0,23.

In gleicher Weise kann der Einfluß einzelner Prädiktoren bzw. Gruppen von Prädiktoren im Vergleich zum vollen Modell interferenzstatistisch abgesichert bzw. in seiner Größenordnung beurteilt werden.[16]

Wie schon ***Urban*** erwähnt hat, werden Logit-Modelle vielfach auch im Sinne von Diskriminanzanalysen eingesetzt. Wir fragen also danach, wie gut durch das Modell die (bekannte) Klassifikation der Fälle hinsichtlich der "abhängigen" Variablen, der staatsanwaltlichen Erledigung, reproduziert werden kann. Hierfür lassen sich verschiedene Regeln formulieren (s. *Anderson/Philips* 1981; als Beispiel vgl. *Ashby et al.* 1986). Folgt man derjenigen, welche zu der größten Zahl von richtigen Zuordnungen führt, so klassifiziert man die Fälle in diejenige Kategorie, für welche sie die höchste Wahrscheinlichkeit aufweisen.[17] Die Wahrscheinlichkeiten werden gemäß Formel (4) berechnet. In dem einfachen Beispiel der Tabelle 1 würden also alle Fälle mit 1, 2 oder 3/4 Delikten in die Gruppe der Anklagen vor dem Einzelrichter, diejenigen mit 5 und mehr Delikten in die Gruppe der Anklagen vor dem Jugendschöffengericht eingeordnet. Hieran läßt sich erkennen, daß die Klassifikation mit durch die Häufigkeit der einzelnen Ausprägungen der abhängigen Variablen bestimmt wird. Tatsächlich zeigt sich auch in dem relativ komplexen Modell aus Tabelle 3, daß kein einziger Fall in die zweite oder dritte Gruppe (also Verfahrenseinstellungen "mit Reaktion" oder Antrag auf vereinfachtes Verfahren) klassifiziert würde. Wir haben die Klassifikation des Modells im Verhältnis zu den beobachteten Werten in Tabelle 4 angegeben. Dort sind zum Vergleich auch die Vorhersagen aus drei weiteren Modellen (mit den gleichen Prädiktoren) enthalten:

- einem *multinomialen Logit-Modell,*
- einem *linearen Regressionsmodell* sowie
- einer *linearen Diskriminanzanalyse.*

Wenn man die individuelle Klassifikation der Fälle durch das *ordinale Logit-Modell* in *Tabelle 4 a)* mit der scheinbar schon relativ guten Reproduktion der Wahrscheinlichkeiten

---

16 Letzteres wäre also die (bei mehreren Prädiktoren: multiple) quadrierte semipartielle Korrelation, vgl. dazu *Andreß* 1986, S. 104 ff. und 112 ff.

17 Dies ist auch die Verfahrensweise von GAUSS und LIMDEP.

durch eine einzige Variable in Tabelle 1 vergleicht, könnte man zunächst die Leistungsfähigkeit des ordinalen Logit-Modells als nicht sehr hoch einzustufen geneigt sein. Tatsächlich führt dieses Modell zu 55,9 % richtiger Klassifikationen, was zwar im Vergleich zu einer rein zufälligen Klassifikation, nicht aber im Vergleich zu einem wünschbaren Ergebnis weitgehend richtiger Einstufungen als gut gelten kann. Vergleicht man jedoch das ordinale Logit-Modell mit möglichen Alternativen, so zeigt sich, daß es die vorhandene Information relativ gut ausnutzen dürfte: Mit einem *multinomialen Logit-Modell* werden zwar einige Fälle in die Gruppen 2 und 3 eingestuft, die Zahl richtiger Klassifikationen liegt allerdings gerade um 11, also 0,2 % der gesamten Stichprobe, über der des ordinalen Logit-Modells. Obwohl das multinomiale Logit-Modell also 48 Koeffizienten mehr schätzt als das ordinale Logit-Modell (und auf diese Weise zu einem Pseudo-R^2 von 0,26 kommt), zahlt sich dies auf der Ebene der individuellen Klassifikation praktisch nicht aus. Noch wesentlich schlechter schneiden die beiden Modelle hinsichtlich der richtigen Zuordnung ab, welche in Tabelle 4c) und 4d) enthalten sind. Das *lineare Regressionsmodell* scheint zwar oberflächlich betrachtet mit dem ordinalen Logit-Modell übereinzustimmen. Der Bereich der durch das lineare Regressionsmodell vorhergesagten Werte für die abhängige Variable liegt jedoch zwischen 1,7 und 6,5! Das Regressionsmodell schätzt also Werte für die abhängige Variable, welche empirisch überhaupt nicht möglich sind. Aber selbst wenn man alle Fälle, die einen vorhergesagten Wert von 5,5 übersteigen, mit 5 codiert (wie in Tab. 4 geschehen), ist die Klassifikationsleistung relativ schlecht, weil kein einziger Fall mit dem Wert 1 richtig vorhergesagt wird. Insgesamt werden durch das lineare Regressionsmodell nur 32,2 % der Fälle richtig vorhergesagt.

Um einiges besser, aber immer noch schlechter als die der beiden Logit-Modelle, sind die Schätzungen anhand einer *linearen Diskriminanzanalyse.* Diese klassifiziert 45,4 % aller Fälle richtig und liegt damit zwischen den beiden Logit-Modellen und dem linearen Regressionsmodell.

Tabelle 4: Beobachtete und durch verschiedene multivariate Modelle vorhergesagte Klassifikation jugendstaatsanwaltlicher Entscheidungen (vgl. Tabelle 3)

a) Ordinales Logit-Modell

	VORHERGESAGT					
BEOBACHTET	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Summe
Gruppe 1	**793**	0	0	412	1	1206
Gruppe 2	440	**0**	0	215	1	656
Gruppe 3	274	0	**0**	271	3	548
Gruppe 4	250	0	0	**1895**	142	2287
Gruppe 5	2	0	0	338	**284**	624
Summe	1759	0	0	3131	431	5321

Anzahl richtiger Klassifikationen (in der Tabelle fett): 2972 (55,9%)

b) Multinomiales Logit-Modell

	VORHERGESAGT					
BEOBACHTET	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Summe
Gruppe 1	**876**	10	6	314	0	1206
Gruppe 2	465	**13**	4	173	1	656
Gruppe 3	288	13	**7**	239	1	548
Gruppe 4	331	11	11	**1820**	114	2287
Gruppe 5	5	1	0	351	**267**	624
Summe	1965	48	28	2897	383	5321

Anzahl richtiger Klassifikationen: 2983 (56,1%)

c) Lineares Regressionsmodell

	VORHERGESAGT					
BEOBACHTET	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Summe
Gruppe 1	**0**	757	414	34	1	1206
Gruppe 2	0	**440**	192	24	0	656
Gruppe 3	0	269	**232**	44	3	548
Gruppe 4	0	228	1150	**747**	162	2287
Gruppe 5	0	2	79	246	**297**	624
Summe	0	1696	2067	1095	463	5321

Anzahl richtiger Klassifikationen: 1716 (32,2%)

d) Lineare Diskriminanzanalyse

	VORHERGESAGT					
BEOBACHTET	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Summe
Gruppe 1	**499**	387	165	145	10	1206
Gruppe 2	193	**278**	118	58	9	656
Gruppe 3	113	164	**177**	79	15	548
Gruppe 4	257	215	354	**1006**	455	2287
Gruppe 5	8	3	18	143	**452**	624
Summe	1070	1047	832	1431	941	5321

Anzahl richtiger Klassifikationen: 2412 (45,4%)

Allerdings ändert sich die Bewertung der Klassifikationsleistung der verschiedenen Modelle, wenn man sich nicht nur auf die völlig korrekten Klassifikationen bezieht, sondern z.B. danach fragt, wieviel Fälle in die richtige Kategorie oder aber in *eine unmittelbar neben der richtigen Kategorie liegende Gruppe* eingeordnet werden, was angesichts der ordinalen Natur der abhängigen Variablen u.U. als sinnvoll angesehen werden kann (vgl. ***Ashby et al.*** 1986). Die vier Modelle sagen in diesem Sinne als richtig oder "beinahe richtig" voraus:

Ordinales Logit:	4163	(78,2%)
Multinomiales Logit:	4190	(78,7%)
Lineare Regression:	4536	(85,2%)
Diskriminanzanalyse:	4305	(80,9%)

Hier schneidet also das lineare Regressionsmodell am besten ab, weil es in ganz erheblichem Umfang Fälle als "beinahe richtig" klassifiziert, während die beiden Logit-Modelle, indem sie die Kategorien 2 und 3 ganz bzw. weitgehend vermeiden, zu einer relativ geringen Anzahl solcher Klassifikationen kommen.

In ähnlicher Weise drückt sich dies aus, wenn wir in Analogie zum Vorgehen bei ***Urban*** (1990) den Zusammenhang zwischen beobachteten und vorhergesagten Werten in einem Korrelationsmaß ausdrücken. Wir betrachten dazu einmal den Koeffizienten Gamma, welcher auch als PRE-Maß interpretiert werden kann (vgl. ***Benninghaus*** 1979), zum anderen Spearman's Rho.[18] Dieser Koeffizient ist zwar im Grunde genommen bei so vielen "Ties" wenig sinnvoll, jedoch berücksichtigt er als einziger den Abstand zwischen den verschiedenen Ausprägungen der Variablen *(Tabelle 5).*

Tabelle 5: Zusammenhang zwischen beobachteter und vorhergesagter Klassifikation für verschiedene Modelle (vgl. Tab. 4):

	Gamma	Spearman's Rho
Ordinales Logit	0,79	0,62
Multinomiales Logit	0,80	0,63
Lineare Regression	0,77	0,66
Lineare Diskriminanzanalyse	0,68	0,63

[18] Dieser Koeffizient wird von *McKelvey/Zavoina* 1975 vorgeschlagen.

Während also bei Betrachtung von Gamma, welches nicht berücksichtigt, wie weit die Schätzungen des Modells von den beobachteten Werten entfernt liegen, das ordinale und das multinomiale Logit-Modell am besten abschneiden (überraschenderweise die lineare Diskriminanzanalyse am schlechtesten), kommt in den Werten von Rho zum Ausdruck, daß beim linearen Regressionsmodell die Zahl der "beinahe richtig" klassifizierten Fälle im Vergleich der vier Modelle am größten ist.

Es sollte allerdings beachtet werden, daß es Klassifikationsregeln gibt, welche auch für das ordinale Logit-Modell zwar zu einer geringeren Zahl richtiger, möglicherweise aber einer höheren Zahl "beinahe richtiger" Klassifikationen führen. Eine solche Vorgehensweise besteht darin, für jeden Fall den durch das Modell geschätzten Wert $z = \Sigma\beta_p x_p$ (p = 1 ...k) zu berechnen. Liegt z unter α_1, wird der Fall in die erste Kategorie der abhängigen Variablen eingestuft, für $\alpha_1 \leq z < \alpha_2$ in die zweite Kategorie, usw. (vgl. *Anderson/Philips* 1981, S. 26). In unserem Beispiel ergibt sich damit die Klassifikation in Tabelle 6.

Tabelle 6: Beobachtete und durch ordinales Logit-Modell vorhergesagte Klassifikation staatsanwaltlicher Entscheidungen (alternative Klassifikationsregel)

	VORHERGESAGT					
BEOBACHTET	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4	Gruppe 5	Summe
Gruppe 1	**200**	604	220	182	0	1206
Gruppe 2	113	**334**	107	101	1	656
Gruppe 3	58	221	**113**	153	3	548
Gruppe 4	18	244	362	**1526**	137	2287
Gruppe 5	0	3	8	332	**281**	624
Summe	389	1406	810	2294	422	5321

Nach dieser Regel werden zwar nur mehr 2454 (46,1 %) Fälle in die richtige Kategorie eingeordnet, die Zahl von Einordnungen in die richtige oder eine benachbarte Kategorie beträgt jetzt aber 4483 (84,3 %), erreicht also fast diejenige des linearen Regressionsmodells. Dementsprechend geht der Wert von Gamma für Tabelle 6 auf 0,74 zurück, derjenige von Rho steigt dagegen auf 0,66, also den Wert für das lineare Regessionsmodell.

Nach der Klassifikationsleistung der verschiedenen Verfahren (und auch nach den Signifikanztests für die einzelnen Koeffizienten, welche, wie erwähnt, für das lineare Regressionsmodell im konkreten Fall praktisch zu den gleichen Entscheidungen führten wie für

das ordinale Logit-Modell) könnte man zu dem Schluß kommen, daß die Unterschiede zwischen ordinalem Logit-Modell und linearem Regressionsmodell in der Praxis nicht sehr erheblich sind. Nach unseren Erfahrungen dürfte das zwar in den meisten Fällen zutreffen. Es darf aber nicht vergessen werden, daß grundsätzlich das lineare Regressionsmodell zu falschen Entscheidungen über die Signifikanz einzelner Prädiktoren führen kann, weil es von Voraussetzungen ausgeht, welche bei einer ordinalskalierten Variablen häufig nicht gegeben sind.[19] Auch der starke "Mittelwert-Bias" des linearen Regressionsmodells, welcher in Tabelle 4c) deutlich wird, läßt es ratsam erscheinen, die Anwendung linearer Regressionsmodelle auf Datenstrukturen wie die vorliegende eher mit Vorsicht zu handhaben.

3 Mögliche alternative Verfahren und Ausklang

Das hier vorgestellte ordinale Logit-Modell ist keineswegs die einzige Möglichkeit einer angemessenen multivariaten Modellierung der Einflüsse auf eine ordinalskalierte abhängig ge Variable. Auf folgende mögliche Alternativen sei hier kurz hingewiesen:

• Anstelle der kumulativen Logits ließen sich auch "kumulative Probits" verwenden, so wie auch die Einflüsse auf binäre abhängige Variablen nicht nur mit Logit-, sondern auch mit Probit-Modellen geschätzt werden können. Allerdings unterscheiden sich die Ergebnisse beider Modelle in der Regel nicht, da die Dichtefunktion der logistischen Verteilung und der Standardnormalverteilung recht ähnlich sind. Im übrigen können mit den im Anhang vorgestellten Programmen GAUSS und LIMDEP gleichermaßen Logit- wie Probit-Modelle geschätzt werden.

• Bei *Agresti* (1983, 1984) werden log-lineare Modelle für ordinalskalierte Variablen diskutiert, welche auch mit der SPSS-Prozedur LOGLINEAR geschätzt werden können (s. auch Hinweise im Anhang). Diese Modelle führen in der Regel zu ähnlichen Entscheidungen über die Signifikanz einzelner Parameter. Allerdings berücksichtigen sie nicht die kumulativen, sondern jeweils nur die "benachbarten" Logits.

• Wenn man ordinalskalierte Variablen als unvollkommene Messung von zugrundeliegenden (latenten) intervallskalierten Variablen auffaßt, läßt sich nach den Ergebnissen von *Jöreskog/Sörbom* (1988) die polychorische Korrelation als beste Schätzung für die zugrundeliegende Produkt-Moment-Korrelation in einem LISREL-Modell verwenden

---

19 Im hier verwendeten Beispiel ist z.B. die bei Tests unterstellte Normalverteilungsannahme der Residuen mit Sicherheit nicht gegeben. - Ein Beispiel für Divergenzen zwischen einem linearen Regressionsmodell und einem Probit-Modell (vgl. Abschnitt 3) für ordinale Variablen findet sich bei *WinshipIMare* 1984.

(s. ***GrafflSchmidt*** 1985; ***Pfeifer/Schmidt*** 1987). Diese Vorgehensweise dürfte vor allem dann sinnvoll sein, wenn für die latenten Variablen mehrere Indikatoren zur Verfügung stehen. Alternative Modelle mit ähnlichem Grundgedanken sind z.B. von ***Muthén*** (1983, 1984) oder ***Bye*** *et al.* (1985) entwickelt worden, das LISREL-Modell hat jedoch den Vorteil, in einem standardisierten Programmpaket zur Verfügung zu stehen.

• Schließlich wird in der Literatur auch vorgeschlagen, ordinalskalierte Variablen so zu transformieren, daß sie den Verteilungsannahmen linearer Regressionsmodelle (oder gegebenenfalls anderer gewünschter Verfahren) so gut als möglich entsprechen; gegebenenfalls könnten dann robuste Verfahren zur Modellierung der Zusammenhänge angewendet werden (s. ***Brockett*** 1981; ***Golden/Brockett*** 1987).

Die Vielzahl (wenn auch nicht immer leicht) verfügbarer multivariater Modelle für ordinalskalierte Variablen dürfte wohl die soziologische Debatte der 70er Jahre,[20] ob Verfahren, welche für intervallskalierte Variablen entwickelt wurden, auch bei ordinalskalierten Variablen angewendet werden können, endgültig in den Bereich des Historischen verweisen. Nicht historisch dürfte jedoch die soziologische Konfusion hinsichtlich der Frage sein, ob und unter welchen Umständen ein höheres Meßniveau als das einer Ordinalskala erreicht werden kann. So ist die häufig vertretene Ansicht, daß es sich bei den üblichen Likert-skalierten Items zur Erfassung von Attitüden um Messungen auf Ordinalskalenniveau handelt, keinesfalls unumstritten; es dürfte sinnvoller sein, solche Items als *ungenaue* Messungen auf Intervallskalenniveau zu behandeln.[21] Ebenso sollten Soziologen die Hinweise aus der psychometrischen Literatur zur Kenntnis nehmen, wie sich gerade in diesem Bereich (der Erhebung von Attitüden) die Genauigkeit der Messungen durch entsprechende Verfahren deutlich verbessern läßt. Die Anwendung solcher Verfahren dort, wo dies angemessen ist, würde daneben immer noch eine Vielzahl ordinalskalierter Variablen übriglassen, für die das hier vorgestellte Verfahren eine Analysemöglichkeit darstellt.

LITERATUR

Agresti, A. (1983)
A Survey of Strategies for Modeling Cross-Classifications Having Ordinal Variables.
In: Journal of the American Statistical Association 78: 184-198.

---

[20] Vgl. dazu exemplarisch die Arbeiten von *Kim* 1975, *O'Brien* 1979, *Bollen/Barb* 1981, *Henry* 1982, welche auch jeweils Verweise auf die vorangegangene Diskussion enthalten.

[21] In dieser Hinsicht wird hier der Standpunkt von *Borgatta/Bohrnstedt* 1972 vertreten. Nicht zutreffend dürfte dagegen deren Vermutung sein, daß solche Messungen auch immer näherungsweise einer Normalverteilung folgen.

Agresti, A. (1984)
Analysis of Ordinal Categorical Data.
New York: Wiley.

Albrecht, P.-A. (Hrsg.) (1990)
Informalisierung des Rechts.
Berlin: de Gruyter (im Erscheinen).

Anderson, J. A./Philips, P. R. (1981)
Regression, Discrimination and Measurement Models for Ordered Categorical Variables.
In: Applied Statistics 30: 22-31.

Andreß, H.-J. (1986)
GLIM.
Braunschweig: Vieweg.

Ashby, D./Pocock, S. J./Shaper, A. G. (1986)
Ordered Polytomous Regression: An Example Relating Serum Biochemistry and Haematology to Alcohol Consumption.
In: Applied Statistics 35: 289-301.

Benninghaus, H. (1979, 3. Aufl.)
Deskriptive Statistik.
Stuttgart: Teubner.

Bollen, K. A./Barb, K. H. (1981)
Pearson's r and Coarsely Categorized Measures.
In: American Sociological Review 46: 232-239.

Borgatta, E. F./Bohrnstedt, G. W. (1972)
How One Normally Constructs Good Measures.
In: Sociological Methods and Research 1: 3-12.

Brockett, P. L. (1981)
A Note on Numerical Assignment of Scores to Ranked Categorical Data.
In: Journal of Mathematical Sociology 8: 91-101.

Bye, B. VVGallichio, S. J./Dykacz, J. M. (1985)
Multiple-Indicator, Multiple-Cause Models for a Single Latent Variable with Ordinal Indicators.
In: Sociological Methods and Research 13: 487-509.

Geißler, R./Marißen, N. (1988)
Junge Frauen und Männer vor Gericht.
In: KZfSS 40, S. 505-526.

Golden, L. L./Brockett, P. L. (1987)
The Effect of Alternative Scoring Methods on the Analysis of Rank Order Categorical Data.
In: Journal of Mathematical Sociology 12: 383-419.

Graff, J./Schmidt, P. (1985)
Structural Equation Models for Qualitative Observed Variables.
In: P. Nijkamp (Hrsg.), Measuring the Unmeasurable.
Den Haag: Nijhus.

Guttman, L. (1977)
What is Not What in Statistics.
In: The Staüstician 26: 81-107.

Heidenreich, K. (1987)
Grundbegriffe der Meß- und Testfheorie.
In: E. Roth (Hrsg.), Sozialwissenschaftliche Methoden
München, Wien: Oldenbourg.

Henry, F. (1982)
Multivariate Analysis and Ordinal Data.
In: American Sociological Review 47: 299-304.

Jöreskog, K. G./Sörbom, D. (1988)
LISREL 7. A Guide to the Program and Applications.
Chicago: SPSS, Inc.

Kim, J.-O. (1975)
Multivariate Analysis of Ordinal Variables.
In: American Journal of Sociology 81: 261-298.

Kriz, J. (1973)
Statistik in den Sozialwissenschaften.
Reinbek bei Hamburg: Rowohlt.

Kühnel, S./Jagodzinski, W./Terwey, M. (1989)
Teilnehmen oder Boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSSx.
In: ZA-Information 25: 44-75.

Ludwig-Mayerhofer, W. (1990)
Die staatsanwaltliche Diversionspraxis im Jugendstrafrecht.
In: Albrecht, P.-A. (Hrsg.) (1990): Informalisierung des Rechts.
Berlin: de Gruyter.

Maddala, G. S. (1983)
Limited-dependent and Qualitative Variables in Econometrics.
Cambridge: Cambridge University Press.

McCullagh, P. (1980)
Regression Models for Ordinal Data
In: Journal of the Royal Statistical Society, Series B 42: 109-142.

McKelvey, R./Zavoina, W. (1975)
A Statistical Model for the Analysis of Ordinal Level Dependent Variables.
In: Journal of Mathematical Sociology 4: 103-120.

Muthén, B. (1983)
Latent Variable Structural Equatioa Modeling with Categorical Data.
In: Journal of Econometrics 22: 43-65.

Muthén, B. (1984)
A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators.
In: Psychometrika 49: 115-132.

O'Brien, R. M. (1979)
The Use of Pearson's R with Ordinal Data.
In: American Sociological Review 44: 851-857.

Pfeifer, A./Schmidt, P. (1987)
Die Analyse komplexer Strukturgleichungsmodelle.
Stuttgart: G. Fischer.

Urban, D. (1990)
Multinomiale LOGIT-Modelle zur Bestimmung der Abhängigkeitsstruktur qualitativer Variablen mit mehr als zwei Ausprägungen.
In: ZA-Information 26: 36-61.

Winship, C/Mare, R. D. (1984)
Regression Models with Ordinal Variables.
In: American Sociological Review 49: 512-525.

ANHANG

Programme zur Berechnung ordinaler Logit-Modelle

Die Analysen dieser Arbeit wurden mit GAUSS, Version 2.0 gerechnet. Außer diesem Programm können noch LIMDEP und in der neuesten Version auch BMDP ordinale Logit-Modelle schätzen, ohne daß der Benutzer selbst die Likelihood-Funktion o.a. spezifizieren muß. Mit allen drei Programmen ist auch die Schätzung von binären und multinomialen Logit-Modellen möglich, so daß diese Programme sich für diejenigen anbieten, welche alle Varianten von Logit-Modellen schätzen wollen. Mit SAS können (am Großrechner) ordinale Logit-Modelle berechnet werden, sofern die User Procedure "Logist" verfügbar ist.

Obwohl LIMDEP um einiges komfortabler zu handhaben ist als GAUSS, muß doch auf die erheblichen Unterschiede in der Verarbeitungsgeschwindigkeit zwischen beiden Programmen am PC hingewiesen werden: GAUSS arbeitet - solange man nicht innerhalb der Statistikprozeduren Datentransformationen vornimmt - bei großen Dateien, wie sie hier verwendet wurden, ca. um den Faktor 8 schneller (dies gilt wahrscheinlich für alle Proze-

duren mit ML-Schätzung); aber auch mit Transformationen hat GAUSS die 3- bis 4fache Verarbeitungsgeschwindigkeit.

Trotzdem hier einige Hinweise zu LIMDEP: In die Liste der Regressoren ist eine Konstante (ONE) aufzunehmen. Der Koeffizient für diese Konstante entspricht dem Logit $\ln((1 - p_1) / p_1)$, also im Absolutbetrag, aber nicht im Vorzeichen der Konstante α_1. Die Koeffizienten, welche mit MU(1) bis MU(r-2) bezeichnet sind, geben die Differenz zwischen den Koeffizienten α_2 bis α_{r-1} und α_1 an. Es empfiehlt sich also, einfach das Vorzeichen des Koeffizienten für "ONE" zu vertauschen; um α_2 bis α_{r-1} zu erhalten, sind jeweils die Koeffizienten MU(1) bis MU(r-2) zu α_1 zu addieren.

Für die Benutzung von GAUSS sei auf folgende Fehlerquellen in der Datei "Ordered.arc" hingewiesen: Wenn das Programm bei Verwendung einer einzigen unabhängigen Variablen mit der Meldung "Index out of ränge" abbricht, muß im Quellcode das Statement (in der von mir benutzten Version in Zeile 450, nach der Ausgabe der Deskriptiv-Statistiken)

```
if maxc(minx[2:rows(minx)] == maxx[2:rows(minx)]) == 1
```
geändert werden zu
```
if maxc(minx[l:rows(minx)] == maxx[l:rows(minx)]) == 1.
```

Sofern Datentransformationen mit "dtran" vorgenommen wurden, kann es zu einer Fehlermeldung ("rows don't match") kommen. Dies kann z.B. behoben werden, indem im "Output-Teil" (Zeile 604) des Quellcodes die Zeile

```
omat = ivlbl~abml[ncon+l:nparm,l]~seml ...
```
geändert wird zu: `omat = xlbl~abml ...`

Mit dem BMDP-Modul "PR", welches ab 1990 verfügbar ist, können sowohl multinomiale als auch ordinale Logit-Modelle geschätzt werden. Bei letzteren können als abhängige Variable auch die sog. "benachbarten" Logits verwendet werden. Außerdem sind in diesem Programm auch Verfahren der Modelldiagnostik enthalten. Ob man in der BMDP-typischen Möglichkeit der schrittweisen Modellsuche einen Vor- oder einen Nachteil sieht, dürfte vom Standpunkt des Benutzers abhängen; in jedem Fall lassen sich auch spezifizierte Modelle testen. Kritisch anzumerken ist, daß am PC durch den beschränkten Arbeitsspeicher - in Abhängigkeit von den Fallzahlen - nicht beliebig große Modelle geschätzt werden können. In den meisten Fällen dürfte hieraus kein Problem entstehen, aber für das Modell in Tabelle 3 dieser Arbeit z.B. ist der verfügbare Arbeitsspeicher bei wei-

tem nicht ausreichend[22]. In inhaltlicher Hinsicht ist zu beachten, daß BMDP die abhängige Variable in der in Fußnote 4 angegebenen umgekehrten Reihenfolge spezifiziert.

Mit SPSS ist dagegen nur die Möglichkeit gegeben, sog. *ordinale log-lineare* Modelle zu schätzen (vgl. *Agresti* 1983, 1984); abgesehen davon, daß diese nur bedingt als Ersatz für ordinale Logit-Modelle fungieren können, sind hier auch am Großrechner sehr schnell die Speichergrenzen überschritten, so daß nur Modelle mit sehr wenigen "unabhängigen" Variablen (welche es bei log-linearen Modellen im Grunde ja nicht gibt) geschätzt werden können.

Als Nachtrag zu den Arbeiten von ***Kühnel et al.*** bzw. ***Urban*** sei - neben dem neuen BMDP-Modul "PR" - auf folgendes hingewiesen: SPSS/PC* enthält ab Version 3.1 eine Prozedur zur Schätzung binärer logistischer Regressionsmodelle, welche die bei ***Kühnel et al.*** geschilderten Umwege überflüssig macht; nach wie vor können jedoch keine multinomialen Logit-Modelle geschätzt werden. Daher sei hier noch das Stand-Alone-Programm KALOS/C erwähnt, welches zur Schätzung binärer und multinomialer Logit-Modelle geeignet ist. Hierbei handelt es sich um eine für PC unter UNIX oder MS-DOS adaptierte Version des Programms KALOS von *Roeding/Küsters/Arminger*, welche von Götz ***Rohwer*** am Hamburger Institut für Sozialforschung entwickelt wurde und über ihn bezogen werden kann
(Adresse des Instituts: Mittelweg 36, 2000 Hamburg 13). Dieses sehr einfach zu handhabende Programm dürfte von Interesse für all diejenigen sein, welche mit "ihrem" Statistik-Paket keine multinomialen Logit-Modelle berechnen können, denen aber ein anderes Programm nicht zur Verfügung steht. KALOS/C ist auch ca. um den Faktor 2 schneller als LIMDEP. Grundsätzlich sollten diese Hinweise auch verdeutlichen, daß es "das" Statistik-Paket für Sozialwissenschaftler nicht gibt und daß es sinnvoller sein dürfte, sich in mehrere Programme einzuarbeiten, als auf ein Programm zu warten, welches alle Wünsche erfüllen kann.

Wolfgang Ludwig-Mayerhofer
Universität Bielefeld, Sfb 227, Teilprojekt Cl
Postfach 8640
4800 Bielefeld 1

---

[72] Eine angekündigte Extended-Memory-Version soll diese Probleme jedoch beheben.

Lassen sich mit SPSS^x-Matrix anwenderspezifische Analyseprobleme lösen?
Ein Anwendungstest am Beispiel der multinomialen logistischen Regression[1]

von Steffen M. Kühnel

Seit Version 4.0 ist in SPSS^X mit der Prozedur MATRIX die Möglichkeit gegeben, Analysemodelle über Matrizen-Operationen selber zu programmieren. Im Prinzip sollte es somit auch möglich sein, innerhalb des SPSS^x-Systems statistische Modelle anzuwenden, für die noch keine Standard-Prozedur existiert. Für eine übej didaktische Zwecke hinausgehende Nutzung ist es wichtig, daß solche "selbstgestrickten" Analyse-Programme auch größere Datenmengen in akzeptabler Zeit verarbeiten können.

Ob die SPSS^x-Prozedur MATRIX diesem Anspruch genügen kann, soll in diesem Beitrag an Hand eines Beispiels untersucht werden. Mein Anwendungsbeispiel ist die Implementation des Newton-Raphson-Algorithmus zur Maximum-Likelihood-Schätzung der Regressionskoeffizienten der multinomialen logistischen Regression. Das statistische Modell der logistischen Regression ist in den letzten Ausgaben der ZA-Information in mehreren Beiträgen behandelt worden (vgl. *Ludwig-Mayerhof er,* 1990; *Kühnel u.a.,* 1989; *Urban,* 1990). SPSS^x-Benutzer können jedoch bislang nur den Spezialfall der binären logistischen Regression anwenden. Eine Implementation des allgemeineren multinomialen Modells über die Prozedur MATRIX könnte den Anwenderkreis dieser Modelle möglicherweise erweitern.

Der vorüegende Beitrag ist so aufgebaut, daß die drei Hauptteile auch unabhängig voneinander gelesen werden können. Im ersten Teil gehe ich noch einmal knapp und recht formal auf das statistische Modell der logistischen Regression ein und beschreibe den verwendeten Algorithmus zur Schätzung der Modellparameter. Im zweiten Abschnitt wird über die Erfahrungen bei der Implementation des Modells berichtet.[2] Dabei gehe ich insbesondere auf Probleme ein, die möglicherweise auch bei anderen Anwendungsversuchen auftreten können. Im dritten Teil wird die praktische Anwendbarkeit des Matrixprogramms an einer sozialwissenschaftlichen Fragestellung geprüft. Inhaltlich geht es hier um die Fra-

---

[1] Ich danke *Michael Terwey,* daß er das hier beschriebene Matrixprogramm aus der Sicht eines externen Anwenders testete und wertvolle Hinweise zur Verbesserung gab.

[2] Das resultierende Matrixprogramm ist in ein SPSS^x-Makro eingebettet, das Interessenten auf Anforderung zur Verfügung gestellt wird. In Anhang 1 wird der Aufruf dieses Makros beschrieben.

ge, ob entsprechend der Handlungstheorie von ***Fishbein und Ajzen*** die Handlungsabsicht der beste und letzlich einzige Prädiktor von Handlungen ist.

1. Der Newton-Raphson-Algorithmus zur ML-Schätzung der Parameter eines multinomialen logistischen Regressionsmodells

Die logistische Regression ist ein Analysemodell zur Untersuchung des Einflusses einer oder mehrerer (metrischer) erklärender Variablen auf eine abhängige nominalskalierte Variable. Der Zusammenhang zwischen der abhängigen Variable und den erklärenden Variablen wird dabei so modelliert, daß das Verhältnis der Auftretenswahrscheinlichkeiten von jeweils zwei verschiedenen Kategorien der abhängigen Variable exponential mit einer Linearkombination der erklärenden Variablen variiert. $X_1, X_2, ..., X_K$ seien die insgesamt K erklärenden Variablen und Y die abhängige Variable mit den Kategorien i = 1, 2, ..., I. Für jeden Fall t einer Stichprobe von n Fällen wird im logistischen Regressionsmodell also unterstellt, daß für zwei Kategorien i und j von Y gilt:

(1)
$$\frac{P(y_t=i)}{P(y_t=j)} = \exp\left(\sum_{k=1}^{K} \beta_{ijk} * x_{kt}\right) = e^{x_t'\beta_{ij}}$$

Von den nach Gleichung (1) theoretisch möglichen K*I*(I-1) Regressionskoeffizienten sind tatsächlich nur K*(I-1) nicht redundant. Vertauscht man nämlich in Gleichung (1) Nenner und Zähler, wird deutlich, daß gilt:

(2)
$$\beta_{ijk} = -\beta_{jik}$$

Darüber hinaus gilt auch für beliebige Kategorien i,j und r der abhängigen Variable:

(3)
$$\beta_{ijk} = \beta_{irk} - \beta_{jrk}$$

Letzteres folgt aus:

$$(4) \qquad \frac{P(Y_t=i)}{P(Y_t=j)} = \frac{\left(\frac{P(Y_t=i)}{P(Y_t=r)}\right)}{\left(\frac{P(Y_t=j)}{P(Y_t=r)}\right)}$$

Bei der Modellschätzung eines logistischen Regressionsmodells werden daher nur die K*(I-1) Regressionskoeffizienten β_{irk} einer Referenzkategorie r der abhängigen Variable geschätzt.[3] Die Schätzung dieser Modellparameter kann nach der Maximum-Likelihood-Methode erfolgen. Dabei werden die Koeffizienten β_{irk} so bestimmt, daß die Wahrscheinlichkeit maximal ist, gerade die in der Stichprobe aufgefundenen Werte der abhängigen Variable zu erhalten. Bei korrekter Modellspezifikation und einer einfachen Zufallsauswahl[4] voneinander unabhängiger Realisationen läßt sich die Likelihood-Funktion folgendermaßen darstellen:

$$(5) \qquad L = \prod_{t=1}^{n} \prod_{i=1}^{I} P(Y_t=i)^{d_{it}}$$

In (5) hat die Indikatorvariable d_{it} den Wert Eins, wenn $p(y_t=i)$, ansonsten den Wert Null. Aus Gleichung 1 und der Summierung aller Wahrscheinlichkeiten zur Summe 1 folgt für die einzelnen Wahrscheinlichkeiten:

$$(6) \qquad P(Y_t=i) = e^{x_t'\beta_{ik}} * P(Y_t=r); \qquad i \neq r$$

$$P(Y_t=r) = \frac{1}{1 + \sum_{i=1 \neq r}^{I} e^{x_t'\beta_{ik}}}$$

---

[3] Der Index r für die Referenzkategorie wird dabei oft ausgelassen. Bei der Interpretation der Koeffizienten muß natürlich die jeweils gewählte Referenzkategorie berücksichtigt werden.

[4] Die folgende Funktion läßt sich auch bei geschichteten Stichproben anwenden. Bei endogen geschichteten Stichproben, bei denen die Auswahl über Teilstichproben für die einzelnen Kategorien der abhängigen Variable erfolgt, werden die Regressionskonstanten allerdings nicht konsistent geschätzt (vgl. *Maier u. Weiss*, 1990:207ff.).

Das Einsetzen von (6) in (5) zeigt, daß die Likelihood-Funktion tatsächlich eine mathematische Funktion der Regressionskoeffizienten β_{irk} ist. Statt nun direkt das Maximum dieser Funktion zu bestimmen, wird das Negative des natürlichen Logarithmus der Funktion minimiert:

(7) $$-\ln L = -\sum_{i=1}^{n}\sum_{j=1}^{J} d_{ij}\cdot\ln\left(P(y_i=j)\right) = -\sum_{i=1}^{n} d_i'\cdot\ln(p_i)$$

Bei unterstellter korrekter Modellspezifikation sind die so geschätzten Regressionskoeffizienten asymptotisch um die gesuchten Populationswerte multinormalverteilt. Über die ebenfalls schätzbaren Standardfehler der Schätzer können daher asymptotisch gültige Signifikanztests durchgeführt werden, die dem T-Test in der linearen Regression entsprechen. Darüber hinaus lassen sich auch asymptotisch gültige Chiquadrat-Tests von beliebigen Linearkombinationen der Regressionskoeffizienten berechnen, so daß sowohl Kontraste als auch der Einfluß jeder einzelnen erklärenden Variablen über die Gesamtheit der ihr zugeordneten (1-1) Koeffizienten getestet werden können. Schließlich sind auch Chiquadrat-Differenzentests (Likelihood-Ratio-Tests) hierarchisch geschachtelter Regressionsmodelle möglich. Unterscheiden sich nämlich zwei logistische Regressionsmodelle nur dadurch, daß in einem der beiden Modelle einige Regressionskoeffizienten a priori auf Null fixiert sind, so ist 2 mal die Differenz der Maxima der beiden negativen logarithmierten Likelihood-Funktionen (7) asymptotisch chiquadratverteilt. Den Chiquadrat-Differenzentests entsprechen in der linearen Regression die F-Tests von R^2 bzw. Änderungen von R^2. Ein dem Determinationskoeffizienten R^2 in gewisser Hinsicht analoges PRE-Maß zur Stärke des Zusammenhangs läßt sich durch den Vergleich der Minima von Gleichung (7) bei zwei hierarchisch geschachtelten Modellen konstruieren. Wenn $-\ln L_0$ das Minimum der Gleichung (7) eines Modells ist, das nur Regressionskonstanten enthält, und $-\ln L_1$ das Minimum eines Modells, das zusätzlich erklärende Variablen aufweist, berechnet sich dieses Pseudo-R^2 nach:

(8) $$P^2 = 1 - \frac{-\ln L_1}{-\ln L_0}$$

Die rechnerische Bestimmung der Maximum-Likelihood-Lösung kann nach dem Newton-Raphson-Algorithmus erfolgen.[5] Dazu benötigt man die ersten und zweiten Ableitungen der Funktion (7) nach den unbekannten Parameterwerten. Wenn I die Referenzkategorie ist und $\mathbf{B} = (\beta_{1I}, \beta_{2I}, \ldots, \beta_{I-1,I})$ die Matrix der Regressionskoeffizient $\mathbf{d}_t^*$ die ersten (I-1) Elemente des Vektors $\mathbf{d}_t$ aus Gleichung (7) enthält und $\mathbf{p}_t^*$ die entsprechenden Elemente von $\mathbf{p}_t$, dann ergeben sich die ersten Ableitungen als:

$$\frac{\partial \ln L}{\partial \boldsymbol{B}} = \sum_{t=1}^{n} \boldsymbol{x}_t (\boldsymbol{d}_t^* - \boldsymbol{p}_t^*)' \tag{9}$$

Die Matrix der zweiten Ableitungen berechnet sich nach:

$$\frac{\partial^2 \ln L}{\partial \mathrm{vec}(\boldsymbol{B}) \partial \mathrm{vec}(\boldsymbol{B})'} = -\sum_{t=1}^{n} (\mathrm{diag}(\boldsymbol{p}_t^*) - \boldsymbol{p}_t^* \boldsymbol{p}_t^{*\prime}) \otimes (\boldsymbol{x}_t \boldsymbol{x}_t') \tag{10}$$

In Gleichung (10) steht "vec(.)" für die Umgruppierung der Elemente einer Matrix in einen Spaltenvektor und "diag(.)" für eine Diagonalmatrix, deren Diagonale die Elemente des in der Klammer aufgeführten Vektors enthalten. Das Symbol ⊗ steht für die Bildung des Kronekerprodukts zweier Matrizen.

Der Algorithmus führt zu einer schrittweisen Minimierung von Gleichung (7). Dabei werden in jedem Iterationsschritt s+1 die Koeffizienten des vorherigen Schrittes s korrigiert. Die Korrekturgröße ist dabei das Negative des Produkts der Inversen der Matrix der zweiten Ableitungen mit dem Vektor der ersten Ableitungen. Es gilt also:

$$\mathrm{vec}(\boldsymbol{B}^{s+1}) = \mathrm{vec}(\boldsymbol{B}^{s}) - \left(\frac{\partial^2 \ln L}{\partial \mathrm{vec}(\boldsymbol{B}^{s}) \partial \mathrm{vec}(\boldsymbol{B}^{s})'} \right)^{-1} \left(\frac{\partial \ln L}{\partial \mathrm{vec}(\boldsymbol{B}^{s})} \right) \tag{11}$$

Die endgültige Lösung ist gefunden, wenn in zwei aufeinanderfolgenden Iterationsschritten die Minimierungsfunktion (7) nicht weiter abnimmt bzw. der Vektor der Regressionskoef-

---

[5] Eine auch für Nichtmathematiker verständliche Erklärung der Logik des Newton-Raphson-Algorithmus findet man **in *Maier u. Weiss*** (1990: 84ff.). Meine Darstellung orientiert sich an den Beschreibungen des Algorithmus bei der SAS-Prozedur CATMOD (vgl. *SAS,* 1985: 212f.) und beim Programm KALOS (vgl. ***Röding*** *u.a.* 1985, 5.2-1 ff.).

fizienten sich nicht mehr ändert, da der Vektor der ersten Ableitungen nur Nullen enthält.[6] Als Startwert wird für die Iteration ein Vektor gewählt, bei dem alle Regressionskoeffizienten den Wert Null haben.

Multipliziert man den Wert der Minimierungsfunktion mit 2, ergeben sich die Ausgangsgrößen für die Chiquadrat-Differenzentests. Gleichsam als Nebenprodukt führt der Newton-Raphson-Algorithmus außerdem zu den Schätzungen der Standardfehler und Korrelationen der Schätzer. Die Varianz/Kovarianz-Matrix der Schätzer der Regressionskoeffizienten ist nämlich gerade das Negative der Inversen der Matrix der zweiten Ableitungen aus Gleichung (10).

2. Erfahrungen bei der Programmierung mit SPSS^X-MATRIX

Die SPSS^X-Prozedur MATRIX läßt sich als Interpreter einer spezifischen Programmiersprache verstehen, deren Datenelemente Matrizen sind. Da alle Operationen und Funktionen über Matrizen definiert sind, lassen sich mathematische Formeln sehr schnell in Programmanweisungen umsetzen. Die bekannte Kleinstquadratlösung der Parameterschätzung einer linearen Regressionsfunktion ließe sich beispielsweise durch folgende Anweisung realisieren:

```
COMPUTE B=INV(T(X)*X)*T(X)*Y
```

Neben der sequentiellen Abarbeitung solcher COMPUTE-Anweisungen sind Iterationen über "LOOP ... END LOOP"-Schleifen und bedingte Verzweigungen über "DO IF ... END IF"-Blöcke programmierbar. Direkte Sprunganweisungen und Unterprogrammstrukturen, wie sie etwa aus der Programmiersprache BASIC mit den Anweisungen "GO TO" bzw. "GO SUB" bekannt sind, werden nicht zur Verfügung gestellt.

Mit den Befehlen "GET" bzw. "SAVE" können Daten zwischen SPSS^X-Systemdateien und der Prozedur MATRIX ausgetauscht werden. Dabei ist durch die Angabe der Spezifikation "/MISSING= ..." eine (eingeschränkte) Missing-Data-Behandlung möglich. Die Anweisungen "READ" und "WRITE" ermöglichen den Zugriff auf externe Dateien. Mit der "PRINT'-Anweisung können schließlich Matrizen in die Ausgabedatei eines SPSS^X-Jobs geschrieben werden.

---

[6] Tatsächlich wird die Iteration beendet, wenn der relative Funktionsabfall bzw. die ersten Ableitungen kleiner als eine vorgegebene kritische Größe sind.

Theoretisch erlauben diese Programm-Elemente der Prozedur MATRIX eine schnelle Umsetzung des Newton-Raphson-Algorithmus zur Schätzung der Regressionskoeffizienten der multinomialen logistischen Regression. Tatsächlich lassen sich die im letzten Abschnitt vorgestellten Gleichungen auch mühelos in SPSS^x-MATRIX-Anweisungen umformen. Als recht hinderlich erweist sich dabei allein eine Restriktion der "COMPUTE"-Anweisung: Auf der linken Seite einer Gleichung muß entweder eine vollständige Matrix oder aber ein einziges Element einer Matrix aufgeführt werden. Sollen dagegen nur die Elemente einer Spalte oder einer Zeile einer Matrix angesprochen werden, müssen dazu für jedes Element eigene COMPUTE-Anweisungen ausgeführt werden.

Trotz des raschen Erfolgs bei der Umsetzung des Algorithmus in ein Matrix-Programm konnte der erste Lösungsversuch noch nicht befriedigen. Es stellte sich nämlich heraus, daß das Programm bereits bei relativ kleinen Stichproben (ab etwa 100 Fällen) unverhältnismäßig viel Zeit zur Berechnung der Lösung benötigte. Der Algorithmus wurde daraufhin so umgeschrieben, daß die Berechnung der Minimierungsfunktion und der ersten und zweiten Ableitungen nicht wie in den Gleichungen (7), (9) und (10) fallweise erfolgt, sondern stattdessen durch Matrizen-Operationen über jeweils die gesamte Stichprobe berechnet wird. Dies führt zu einer erheblichen Beschleunigung der Rechenzeit. Offenbar dauert das fallweise Abarbeiten in einer Schleife erheblich länger als die Ausführung einer Operation über Matrizen.

Bei einem solchen stichprobenweisen Vorgehen muß jedoch darauf geachtet werden, daß die einzelnen Matrizen nicht zu groß werden. Zwar kann die SPSS^x-Prozedur MATRIX durchaus Matrizen mit etlichen tausend Elementen bearbeiten. Operationen auf solchen Matrizen können jedoch auch den größten Speicher sprengen. Sind etwa die Elemente jeder Zeile einer Matrix mit einer anderen Zahl zu multiplizieren, kann dies in Matrixschreibweise so formuliert werden, daß die Ursprungsmatrix von links mit einer Diagonalmatrix mit den entsprechenden Faktoren in der Diagonale multipliziert wird. Bei z.B. 3000 Zeilen der Ursprungsmatrix hätte die Diagonal-Matrix entsprechend 3000 mal 3000 Elemente. In solchen Situationen ist es notwendig, vor der Umsetzung von Matrizenoperationen in ein SPSS^x-Matrixprogramm die Operationen umzuformen. Statt also etwa von links mit einer Diagonalmatrix zu multiplizieren, ist alternativ auch eine elementweise Multiplikation mit einer Matrix gleicher Größe wie die Ausgangsmatrix möglich, wobei die Zeilen der Multiplikationsmatrix die entsprechenden Faktoren enthalten. Solche Umformungen verlangen jedoch mehr als rudimentäre Kenntnisse des Matrizenkalküls, was den Kreis der Programmieranwender der Prozedur MATRIX vermutlich stark einschränkt.

Neben der Notwendigkeit, in Matrizenschreibweise dargestellte Algorithmen umzuformulieren, um Verarbeitungszeit zu gewinnen oder Speicherplatz einzusparen, traten bei mei-

nem Versuch, die logistische Regression als SPSS^X-Matrix-Programm zu schreiben, noch einige kleinere Probleme auf, die daher rührten, daß die vorliegende Dokumentation zur Prozedur MATRIX recht knapp gehalten ist und teilweise auch Fehler enthält.[7] So wird etwa in der Kurzbeschreibung der Matrix-Funktion "ALL" behauptet, daß diese Funktion prüft, ob alle Elemente einer Matrix positiv sind. In der detaillierteren Beschreibung der Funktion stellt sich dann heraus, daß diese Funktion tatsächlich prüft, ob alle Elemente einer Matrix ungleich Null sind bzw. mindestens ein Element den Wert Null hat. Ein weiteres Problem trat dadurch auf, daß in der Dokumentation von einer amerikanischen Tastatur ausgegangen wird. Sonderzeichen wie geschweifte Klammern und selbst das Ausrufungszeichen zur Identifikation von Makroschlüsselwörtern mußten auf meiner deutschen Tastatur durch Umlaute dargestellt werden, wobei die "passenden" Umlaute aber zunächst einmal gefunden werden mußten.

Insgesamt traten jedoch keine wirklich schwerwiegenden Probleme bei der Umsetzung des Algorithmus in ein SPSS^X-MATRDC-Programm auf.[8] Das Programm wurde daraufhin noch um einige Zusatzfunktionen erweitert, die zusätzliche Koeffizienten berechnen, Signifikanztests durchführen und deskriptive Statistiken ausgeben. Schließlich wurde um das Matrix-Programm ein SPSS^X-Makro geschrieben, so daß das Programm wie eine SPSS^X-Prozedur aufgerufen werden kann. Als fehlerträchtig zeigte sich hier, daß SPSS^X bei MAKRO- bzw. MATRJX-Anweisungen bisweilen entgegengesetzte Syntaxregeln verwendet. So wird das Ende von Schleifen und bedingten Blöcken in MATRIX-Anweisungen durch "END LOOP" bzw. "END IF" angezeigt, in MAKRO-Befehlen dagegen durch "!DOEND" bzw. "!IFEND". Darüber hinaus scheint es möglicherweise Probleme beim mehrfachen Wechsel zwischen MAKRO- und MATRIX-Anweisungen zum Programmablauf zu geben. Jedenfalls lief mein Programm erst fehlerfrei, nachdem alle "LOOP ... END LOOP"- und "DO IF ... END IF"-Blöcke des MATRIX-Programms innerhalb eines einzigen "!IF ,,. !IFEND "-Blocks der MAKRO-Umgebung untergebracht waren.

Als Ergebnis des Programmierversuchs ist schließlich ein SPSS^X-Makro entstanden, das wie eine standardmäßige SPSS^x-Prozedur mit einem Namen und nachfolgenden Spezifikationen aufgerufen werden kann (siehe Anhang 1). Allerdings ist die Ausgabe dieser selbsterzeugten Prozedur graphisch nicht so schön wie bei den standardmäßigen SPSS^X-Prozeduren. Ursache ist die geringe Mächtigkeit des "PRINT'-Befehls. So können mit einem

---

7 Ich konnte mich bei meinen Versuchen nur auf den "SPSS Reference Guide" stützen. Möglicherweise enthält der "SPSS Advanced Statistics User's Guide" eine umfassendere Dokumentation der SPSS^X-Prozedur MATRIX.

8 Für Interessierte sind die Programmanweisungen zur Parameterschätzung im Anhang 2 wiedergegeben.

"PRINT'-Befehl nicht mehrere Matrizen angesprochen werden. Außerdem kann nur ein einziges Format für alle Elemente der auszugebenen Matrix spezifiziert werden.

Die Größe des Makros wuchs im Laufe der Zeit auf über 500 Zeilen an. Da in der Beschreibung zur SPSS^x-Prozedur MATRIX und zum MAKRO-Interpreter keine Angaben über die maximale Anzahl der Anweisungen innerhalb eines Matrixprogramms bzw. eines Makros gegeben werden, ist unklar, wie groß ein solches Programm maximal sein kann. Auch der Matrix-Befehl "DISPLAY STATUS" war hier nicht hilfreich, zumal mir keine Informationen über die Interpretation der Ausgabe dieses Befehls vorlagen. Unzutreffende Kommentare bei Syntaxfehlern[9] am Ende meiner Programmierversuche weisen darauf hin, daß möglicherweise die Kapazitätsgrenze erreicht wurde. Ähnlich dürfte auch das Phänomen zu deuten sein, wonach unerklärliche Fehlerabbrüche nicht mehr auftraten, wenn vor dem Einbinden und Abarbeiten des Makros mit dem "EXEC'-Befehl Datenmodifikationen ohne nachfolgenden Prozeduraufruf ausgeführt wurden. Überraschenderweise stellte sich auch heraus, daß bei wiederholtem Aufrufen des Makros innerhalb eines SPSS^x-Jobs ab dem dritten Aufruf das Programm mit einer (nicht zutreffenden) Syntaxfehlermeldung abbrach. Eine bessere Dokumentation oder gar eine vom Benutzer beeinflußbare dynamische Alloziierung von Programmspeicherplatz wäre bei der praktischen Arbeit sicherlich sehr hilfreich.

Zur Prüfung der Brauchbarkeit des Programms in (typischen) sozialwissenschaftlichen Analysen wurden zunächst mit den Daten des kumulierten ALLBUS einige logistische Regressionsmodelle geschätzt. Bei diesen Tests traten zunächst neue Probleme auf. Diese betreffen Funktionen der SPSS^x-Prozedur MATRIX. So können einige sehr große Funktionswerte bei der Berechnung der kumulativen Chiquadratverteilung zu Fehlerabbrüchen führen. Weiter zeigte sich, daß die Funktionsberechnung der Determinante einer Matrix gelegentlich mit der Fehlermeldung abbrach, daß die Determinante nicht darstellbar sei. Bei dem Versuch, diesen Fehler dadurch zu umgehen, daß statt der Determinante der Rank einer Matrix berechnet wurde, stellte sich heraus, daß die Matrixfunktion "RANK" offenbar auch nicht immer den tatsächlichen Rank einer Matrix korrekt berechnete. Bei meinem Programm zur logistischen Regression konnten diese Probleme schließlich über die Analyse der Eigenwerte einer Matrix umgangen werden. So führten diese "Kinderkrankheiten" der SPSS^x-Prozedur MATRIX in meinem Anwendungsfall nicht zu falschen Ergebnissen. Jedenfalls waren alle Berechnungen mit der selbstgeschriebenen Prozedur stets identisch mit denen, die zur Kontrolle auch mit der SAS-Prozedur CATMOD berechnet wurden.

---

[9] Recht bizarr war beispielweise eine Fehlermeldung, bei der statt einer Beschreibung des Fehlers der "HELP"-Text der SPSS^x-Prozedur MANOVA ausgegeben wurde.

Bei den Testläufen stellte sich heraus, daß auch große Datensätze verarbeitet werden konnten. So waren durch Heraufsetzen der Obergrenze des virtuellen Arbeitsspeichers auch die Datenmengen des kumulierten ALLBUS mit der SPSS^X-Prozedur MATRIX zu analysieren. Selbst ein Modell mit mehr als 7200 Fällen, 29 erklärenden Variablen und einer abhängigen Variable mit vier Kategorien konnte das Programm bewältigen. Es benötigte allerdings auf dem ZA-Rechner vom Typ IBM 4183 unter CMS 6 relativ viel Zeit zur Berechnung der Lösung.[10] Insgesamt ist das Ergebniss der Programmierung positiv zu bewerten: Nach der Beseitigung der Anfangsprobleme steht nun ein prozedurartiges SPSS^X-Makro zur Verfügung, das allgemein eingesetzt werden kann.

3. Eine realistische Anwendung: Die Prüfung einer Hypothese der Theorie von *Fishbein und Ajzen*

Eine der prominentesten Handlungstheorien der Sozialpsychologie ist die "Theorie bedachter Handlungen" (theory of reasoned action) von *Fishbein und Ajzen* (1975; *Ajzen und Fishbein,* 1980; *Ajzen,* 1988). Nach dieser Theorie sind Handlungen als Realisationen von Handlungsabsichten aufzufassen, welche wiederum durch die Einstellung zum Verhalten und die subjektive Norm bezüglich des Verhaltens erklärt werden. Andere Variablen können als externe Faktoren allein indirekt über die Beinflussung der Einstellung, der subjektiven Norm oder des relativen Gewichts von Einstellung respektive Norm auf die Handlungsabsicht und deren Ausführung wirken.

Mit den Daten, die im Rahmen der Begleituntersuchung zur Volkszählung 1987 erhoben wurden (vgl. *Scheuch u.a.,* 1989), soll im folgenden die Hypothese geprüft werden, daß allein die Verhaltensabsicht einen signifikanten Einfluß auf das Teilnahmeverhalten bei der Volkszählung 1987 hatte. Insbesondere sollte nach der Theorie bedachter Handlungen die Einstellung zur Volkszählung keinen eigenständigen Einfluß neben der Verhaltensabsicht haben. In der Begleituntersuchung zur Volkszählung 1987 wurde kurz vor dem Beginn der Volkszählung u.a. die Verhaltensabsicht erhoben. Ein gutes halbes Jahr später gaben die Befragten Auskunft über ihr tatsächlich realisiertes Verhalten. Bei der Handlungsabsicht wie beim späteren Verhaltensbericht konnten die Befragten zwischen fünf Antwortkategorien wählen: a) alle Fragen so gut wie möglich beantworten, b) einige Fragen nicht beantworten, c) einige Fragen nicht wahrheitsgemäß beantworten, d) den Fragebogen weitgehend nicht wahrheitsgemäß beantworten und e) bestimmt nicht teilnehmen bzw. den Fragebogen nicht beantworten.

---

[10] Eine genaue Angabe über die benötigte Rechenzeit konnte nicht ermittelt werden, da die vom SPSS^X-System ausgedruckten CPU-Zeiten bei einigen Jobs offensichtlich nicht zutrafen.

In der folgenden Analyse wird die Wahl der ersten Antwortkategorie als *Kooperation*, die der letzten als *offener Boykott* und eine Antwort in einer der drei mittleren Kategorien als *verdeckter Boykott* bezeichnet. Die Zusammenfassung der mittleren Kategorien läßt sich damit rechtfertigen, daß bei diesen Handlungsalternativen die Aufforderung der Zählstellen zur Teilnahme nicht vollständig ignoriert wird. Man kann die Ansicht vertreten, daß die drei Kategorien eine ordinale Variable definieren und entsprechend das logistische Modell für ordinale Daten anwenden, das *Ludwig-Mayerhofer* in seinem Beitrag in diesem Heft vorstellt. Nach den Ergebnissen von anderen Analysen dieser Daten *(Kühnel u. Scheuch,* 1990) scheint die Auffassung einer nominalskalierten Variablen allerdings angemessener zu sein.

Die affektiv-emotionale Einstellung zur Volkszählung wurde in der Begleituntersuchung durch die Frage nach der generellen Haltung zur Zählung erfaßt, wobei sieben Antwortmöglichkeiten von *"sehr ablehnend"* (1) bis *"sehr zustimmend"* (7) vorgegeben waren. Nach der Theorie von *Fishbein und Ajzen* sollte die vor der Volkszählung erhobene Einstellung nach der Kontrolle der Verhaltensabsicht keinen Effekt auf das (berichtete) Teilnahmeverhalten haben. Zur Prüfung dieser Hypothese habe ich mit meinem Matrixprogramm ein logistisches Regressionsmodell analysiert, in dem das nach der Zählung berichtete Teilnahmeverhalten durch die vor der Zählung erhobene Einstellung zur Volkszählung und die Teilnahmeabsicht erklärt wird. Da die Verhaltensabsicht wie das berichtete Antwortverhalten als eine trichotome nominalskalierte Variable aufgefaßt wird, ist diese erklärende Variable in zwei dichotome Dummy-Variablen aufgespaltet worden, die den Wert Eins annehmen, wenn ein Befragter als Verhaltensabsicht offenen Boykott (OFFEN) bzw. verdeckten Boykott (VERDECKT) angibt. Befragte, die sich kooperativ verhalten wollen, haben bei beiden Variablen den Wert Null. Zusammen mit der Regressionskonstante hat das Regressionsmodell somit insgesamt vier Prädiktoren. Dieses Modell wird durch die SPSS^x-Anweisung:

```
MLR VAR=VZTEILN OFFEN VERDECKT VZEINST1 /RKAT=0
```

berechnet, wobei "VZTEILN" der Variablenname für das berichtete Antwortverhalten ist und "OFFEN", "VERDECKT" und "VZEINST1" die drei erklärenden Variablen sind.[11]

Tabelle 1 zeigt eine Zusammenfassung der vom Programm ausgedruckten Koeffizienten. Da die abhängige Variable drei Kategorien aufweist, sind jedem Prädiktor zwei Regressionskoeffizienten zugeordnet, so daß das Modell insgesamt 8 Parameter aufweist. Der

[11] Eine vollständige Beschreibung des Aufrufs des Matrixprogramms befindet sich in Anhang 1 dieses Beitrags.

Tabelle ist zu entnehmen, daß sich die Devianz, d.h. -2 mal der Wert der logarithmierten Likelihood-Funktion (5) bei 1156 gültigen Fällen von 2539,99 auf 610,04 verringert. Eine ganz erhebliche Devianzreduktion erhält man jedoch auch schon in einem Modell, das nur die Regressionskonstante enthält. Bezogen auf dieses Konstantenmodell verringert sich die Devianz nur um 134,74. Bei sechs Freiheitsgraden (den sechs Regressionskoeffizienten der beiden Dummy-Variablen und der Einstellung zur Volkszählung) ist dieser chiquadratverteilte Wert signifikant von Null verschieden. Anteilsmäßig beträgt die Devianzreduktion gegenüber dem Konstantenmodell 18,1 %. Dieser Prozentwert kann als Analogon zum Determinationskoeffizienten in einem linearen Regressionsmodell aufgefaßt werden.

Da zu jedem Prädiktor zwei Regressionskoeffizienten gehören, ergibt sich für den Signifikanztest der Prädiktoren eine Chiquadratverteilung mit zwei Freiheitsgraden. Die Nullhypothese postuliert, daß beide Koeffizienten Null sind. Obwohl die Werte der Regressionskoeffizienten von der gewählten Referenzkategorie abhängen, führt der Test bei beliebigen Referenzkategorien zum gleichen Ergebnis. Tabelle 1 ist zu entnehmen, daß alle Prädiktoren einen signifikanten Einfluß auf die abhängige Variable haben. Inhaltlich bedeutet dies insbesondere, daß auch die Einstellung zur Volkszählung neben der Verhaltensabsicht einen eigenen direkten Einfluß auf das (berichtete) Teilnahmeverhalten hat, die Hypothese aus der Theorie von *Fishbein und Ajzen* demnach falsch zu sein scheint.

Der Blick auf die Regressionskoeffizienten zeigt, in welche Richtung die erklärenden Variablen wirken. Wird die Kategorie *"Kooperation"* als Referenzkategorie gewählt, so sieht man am negativen Vorzeichen der Regressionskoeffizienten, daß sich bei einem höheren Wert auf der Variable "VZ-Einstellung" das Wahrscheinlichkeitsverhältnis zuungunsten der Antwortkategorien *"verdeckter Boykott"* (ß=-0.421) bzw. *"offener Boykott"* (ß=-0.477) entwickelt. Aus den T-Werten, die sich aus den Quotienten der Regressionskoeffizienten, geteilt durch die zugehörigen Standardfehler, ergeben, wird ersichtlich, daß beide Koeffizienten auf dem 5%-Niveau signifikant von Null verschieden sind. Die Exponentation der Regressionskoeffizienten wird in der Tabelle 1 als Effekt bezeichnet (vgl. *Long*, 1987). Die Werte geben an, um welchen Faktor sich das Wahrscheinlichkeitsverhältnis ändert, wenn die erklärende Variable um +1 Einheit zunimmt. Zum Vergleich der relativen Einflüsse der Variablen geben die standardisierten Effekte an, wie sich die Wahrscheinlichkeitsverhältnisse bei einer Änderung eines Prädiktor um +1 Standardabweichung ändern. Bei Werten kleiner Eins ist es sinnvoll, den Kehrwert anzugeben.

Der Tabelle ist zu entnehmen, daß die Chancen der *"Kooperation"* gegenüber dem *"verdeckten Boykott"* um etwa das Eineinhalbfache (1.52) steigen, wenn die Einstellung um +1 Einheit zunimmt. Gegenüber dem *"offenen Boykott"* steigt die Chance der *"Kooperation"* sogar noch etwas stärker (um den Faktor 1,61). Inhaltlich bedeutet dies, daß mit zuneh-

Tabelle 1: Logistische Regression des berichteten VZ-Teünahmeverhaltens auf die Teilnahmeabsicht und die Einstellung zur Volkszählung (vor der Zählung)

Modellanpassung:

Modell	Parameter zahl	Devianz
Nullmodell	0	2539.99
Konstantenmodell	2	744.78
Endmodell	8	610.04

Devianzreduktion: 18.1 % (Chiquadrat: 134.74, df: 6, P < 0.0005)

Fallzahl: 1156

Signifikanztest der Prädiktoren:

Variable	Chiquadrat	df	P
Offen	19.069	2	<.0005
Verdeckt	21.096	2	<.0005
VZ-Einstellung	29.830	2	<.0005
(Konstante)	34.041	2	<.0005

Regressionskoeffizienten:

a) Verdeckter Boykott zu Kooperation

Prädiktor	Beta	Std. Fehler	T-Wert	Effekt	Std.Eff.
Offen	1.589	.505	3.15	4.90	1.31
Verdeckt	1.219	.293	4.16	3.38	1.78
VZ-Einstellung	-0.421	.082	-5.15	1/1.52	1/2.31
(Konstante)	-1.798	.380	-4.74	1/6.04	

b) Offener Boykott zu Kooperation

Prädiktor	Beta	Std.Fehler	T-Wert	Effekt	Std.Eff.
Offen	4.004	1.209	3.31	54.83	1.99
Verdeckt	2.212	1.0 90	2.03	9.13	2.84
VZ-Einstellung	-0.477	.231	-2.06	1/1.61	1/2.58
(Konstante)	-4.548	1.276	-3.56	1/94.40	

c) Verdeckter Boykott zu offenem Boykott

Prädiktor	Beta	Std.Fehler	T-Wert	Effekt	Std.Eff.
Offen	-2.415	1.266	-1.91	1/11.19	1/1.51
Verdeckt	-0.993	1.123	-0.88	1/2.70	1/1.60
VZ-Einstellung	0.056	.242	0.23	1.06	1.12
(Konstante)	2.749	1.319	2.08	15.63	

mender positiver Einstellung zur Volkszählung die Wahrscheinlichkeit der *"Kooperation"* gegenüber dem *"verdeckten Boykott"* selbst dann steigt, wenn die Teilnahmeabsicht als Kontrollvariable im Regressionsmodell berücksichtigt wird.

Die unterste Teiltabelle zeigt den Einfluß der Prädiktoren, wenn die Kategorie *"offener Boykott"* als Referenzkategorie gewählt wird. Entsprechend Gleichung (3) ergeben sich die Regressionskoeffizienten aus der Differenz der beiden anderen Tabellen. Zur Berechnung der Standardfehler und der T-Werte benötigt man die Kovarianzen bzw. Korrelationen der Schätzer. Das Matrixprogramm zur logistischen Regression berechnet durch die Angabe der Spezifikation "/RKAT=0" Regressionskoeffizienten, Standardfehler, T-Werte, Wahrscheinlichkeiten und unstandardisierte und standardisierte Effekte für alle möglichen Referenzkategorien der abhängigen Variable.

Bei der Erläuterung der Koeffizienten aus Tabelle 1 wurde bereits darauf hingewiesen, daß die Ergebnisse dieser Analyse als Falsifikation der Hypothese von ***Fishbein und Ajzen*** aufgefaßt werden können, nach der allein die Verhaltensabsicht die Ausführung einer Handlung beeinflußt. Diese Bewertung kann jedoch etwas vorschnell sein. So weisen ***Fishbein und Ajzen*** (1975:370) selber darauf hin, daß diese Beziehung nur gilt, wenn zwischen der Messung der Verhaltensabsicht und der geplanten Verhaltensausführung keine Ereignisse eintreten, die die Verhaltensabsichten verändern. Tatsächlich haben wir in der Begleituntersuchung zur Volkszählung feststellen können, daß bei Beginn der Zählung eine deutliche Änderung der Einstellung zur Zählung zugunsten einer positiveren Bewertung festzustellen war (vgl. ***Kühnel,*** 1988:36; ***Scheuch*** *u.a.,* 1989:33). Wenn diese Veränderungen die Verhaltensabsichten beeinflußt haben, ist die erhobene Verhaltensabsicht **vor** der Zählung nur ein unvollkommener Schätzer der nach der Theorie letztlich relevanten Verhaltensabsicht **bei** der Zählung. Da andererseits die Einstellung vor der Zählung mit der späteren Einstellung während der Zählung korreliert ist, kann die Einstellung vor der Zählung als eine verzerrte Messung der Einstellung bei der Zählung aufgefaßt werden. Da theoriegemäß die Einstellung bei der Zählung die Verhaltensabsicht bei der Zählung beeinflussen kann, ist bei dieser Interpretation durchaus damit zu rechnen, daß die vor der Zählung erhobene Verhaltensabsicht und die Einstellung zur Zählung beide das spätere Teilnahmeverhalten beeinflussen.

Zur Überprüfung dieser Vermutung habe ich in einem zweiten Analyseschritt die nach der Volkszählung erhobene Einstellung zur Zählung als weiteren Prädiktor zur Erklärung des berichteten Antwortverhaltens herangezogen, da aufgrund der genannten früheren Analysen ***(Kühnel,*** 1988; ***Scheuch u.a.,*** 1989) diese Messung als relativ reliabler Indikator der Einstellung zur Volkszählung während der Zählung aufgefaßt werden kann. Die Ergebnisse dieser zweiten Regressionsschätzung sind in Tabelle 2 festgehalten.

Tabelle 2: Logistische Regression des berichteten VZ-Teilnahmeverhaltens auf die Teilnahmeabsicht und die Einstellung zur Volkszählung (vor und nach der Zählung)

Modellanpassung:

Modell	Parameterzahl	Devianz
Nullmodell	0	2533.40
Konstantenmodell	2	744.26
Endmodell	10	552.16

Devianzreduktion: 25.8 % (Chiquadrat: 192.10, df: 8, P < 0.0005)

Fallzahl: 1153

Signifikanztest der Prädiktoren:

Variable	Chiquadrat	df	P
Offen	14.549	2	.001
Verdeckt	17.510	2	<.0005
Einstellung vor VZ	4.501	2	.105
Einstellung nach VZ	47.877	2	<.0005
(Konstante)	5.622	2	.060

Regressionskoeffizienten:

a) Verdeckter Boykott zu Kooperation

Prädiktor	Beta	Std.Fehler	T-Wert	Effekt	Std.Eff.
Offen	1.523	.548	2.78	4.59	1.30
Verdeckt	1.133	.300	3.78	3.10	1.71
Einst. vor VZ	-0.190	.090	-2.12	1/1.21	1/1.46
Einst. nach VZ	-0.485	.079	-6.14	1/1.62	1/2.40
(Konstante)	-0.530	.425	-1.25	1/1.70	

b) Offener Boykott zu Kooperation

Prädiktor	Beta	Std.Fehler	T-Wert	Effekt	Std.Eff.
Offen	3.870	1.278	3.03	47.97	1.94
Verdeckt	2.131	1.099	1.94	8.43	2.74
Einst. vor VZ	-0.054	.243	-0.22	1/1.06	1/1.11
Einst. nach VZ	-0.840	.224	-3.75	1/2.32	1/4.55
(Konstante)	-2.822	1.348	-2.09	1/16.81	

c) Verdeckter Boykott zu offenem Boykott

Prädiktor	Beta	Std.Fehler	T-Wert	Effekt	Std.Eff.
Offen	-2.347	1.306	-1.80	1/10.46	1/1.50
Verdeckt	-0.999	1.128	-0.89	1/2.71	1/1.60
Einst. vor VZ	-0.136	.253	-0.54	1/1.15	1/1.31
Einst. nach VZ	0.355	.230	-1.54	1.43	1.90
(Konstante)	2.291	1.386	1.65	9.89	

Tatsächlich ergibt diese Analyse, daß nun die Einstellung vor der Zählung keinen signifikanten Einfluß auf das berichtete Teilnahmeverhalten hat. Bei zwei Freiheitsgraden ergibt sich für diesen Prädiktor ein Chiquadratwert von 4,5 (P: 0,105). Betrachtet man jedoch die einzelnen Regressionskoeffizienten, so bleibt doch ein leichter Effekt der Einstellung vor der Zählung auf das Wahrscheinlichkeitsverhältnis von *"verdecktem Boykott"* zu *"Kooperation"* sichtbar, der auch auf dem 5%-Niveau signifikant ist (T-Wert: 2.12).

Insgesamt ergibt sich somit ein Ergebnis, das sowohl zugunsten als auch zuungunsten der Theorie von ***Fishbein und Ajzen*** interpretiert werden kann. Aus einer Perspektive, nach der im Sinne von ***Lakatos*** (1982) Theorien erst aufgegeben werden sollen, wenn adäquatere Theorien formuliert sind, läßt sich das Gesamtergebnis m.E. eher als eine Bestätigung der Theorie von ***Fishbein und Ajzen*** auffassen.[12]

4. Resümee

In diesem Beitrag habe ich an einem Beispiel untersucht, ob es möglich ist, mit der Prozedur MATRIX Analysen durchzuführen, für die im Rahmen von SPSS^X keine Standardprozedur zur Verfügung gestellt wird. Zumindest für das Anwendungsbeispiel läßt sich diese Frage positiv beantworten. Nach anfänglichen Schwierigkeiten war es möglich, ein SPSS^X-Makro zu programmieren, mit dem sich Daten der in der quantitativen Sozialforschung üblichen Größenordnung wie mit einer standardmäßigen SPSS^x-Prozedur analysieren lassen. Die Einbindung der Prozedur MATRIX in das SPSS^x-System kann daher als eine erhebliche Verbesserung dieses Programmpaketes betrachtet werden.

Diese generelle Beurteilung ist allerdings zu relativieren. Zum einen bin ich bereits in der ersten Anwendung der Prozedur MATRIX auf offensichtliche Programmierfehler gestoßen. Es ist zu hoffen, daß solche Fehler schnell gefunden und in den nächsten Versionen behoben werden. Grundsätzlich vermeiden lassen sich bei einem großen Programmsystem Programmierfehler allerdings kaum. Wünschenswert wäre auch eine umfassendere Dokumentation, die insbesondere auf die Grenzen des Programms eingeht.

Vergleicht man die SPSS^x-Prozedur MATRIX mit Konkurrenzprodukten, wie etwa der PROC MATRIX bzw. dem IML-System von SAS, oder mit dem PC-System GAUSS, so erreicht der Matrixinterpreter in SPSS^x noch nicht deren Leistungsumfang. Dies liegt auch daran, daß in die SPSS^x-Prozedur MATRIX nicht alle Funktionen implementiert wurden,

---

[12] Bei einer inhalüich orientierten Arbeit wäre an dieser Stelle eine Einbeziehung weiterer erklärender Variablen notwendig. Eine umfassende Analyse würde jedoch den Rahmen dieses eher methodischen Beitrags sprengen.

die in SPSS^X bei der Aufbereitung der Daten zur Verfügung stehen. Eine ernsthafte Einschränkung der Anwendungsmöglichkeiten sehe ich hierin allerdings nicht.

Anwendungsgrenzen dürften eher darin bestehen, daß die Beherrschung des Matrizenkalküls nicht zu den Techniken gehört, die üblicherweise von empirisch arbeitenden Sozialwissenschaftlern verlangt werden. Soll das mit der Prozedur MATRIX bereitgestellte Instrument über den Kreis der Methodenexperten hinaus genutzt werden, wird es notwendig sein, von Spezialisten geschriebene Matrixprogramme einem breiteren Publikum zur Verfügung zu stellen.

Literatur:

Ajzen, I. (1988), Attitudes, Personality, and Behavior. Stony Stratford: Open University.

Ajzen, I. u. M. Fishbein (1980), Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs: Prentice-Hall.

Fishbein, M. u. I. Ajzen (1975), Belief, Attitüde, Intention and Behavior: An Introduction to Theory and Research. Reading, Mass.: Addison-Wesley.

Ludwig-Mayerhofer, W. (1990), "Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen. ZA-Information 27 (in diesem Heft).

Maier, G. u. P. Weiss (1990), Modelle diskreter Entscheidungen. Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften. Wien: Springer.

Kühnel, SM. (1988), "Full-Quasi-Likelihood-Schätzung: Die Ausnutzung aller verfügbaren Informationen bei der Schätzung von linearen Strukturgleichungsmodellen." ZA-Information 23:24-46

Kühnel, SM., W. Jagodzinski u. M. Terwey (1989), "Teilnehmen oder Boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSS^X." ZA-Information 25: 44-75.

Kühnel, SM. u. E.K. Scheuch (1990), Der Einfluß der Strafandrohung auf das Teilnahmeverhalten bei der Volkszählung. Ergebnisse der Kölner Begleituntersuchung zur Volkszählung. Köln: Unveröffentlichtes Vortragsmanuskript.

Lakatos, I. (1982), "Falsifikation und die Methodologie wissenschaftlicher Forschungsprogramme." S. 7-107 in: /. ***Lakatos,*** Philosophische Schriften, herausgegeben von ***J. Worral und G. Currie,*** Band 1: Die Methodologie der wissenschaftlichen Forschungsprogramme. Braunschweig: Vieweg.

Long, J.S. (1987), "A Graphical Method for the Interpretation of Multinomial Logit Analysis." Sociological Methods and Research, 15:420-446.

Röding, M., U. Küsters u. G. Arminger (1985), KALOS Version 1.0. Ein interaktives Programmsystem zur Analyse kategorialer Logitmodelle. Programmbeschreibung. Köln: ZA.

SAS Institute Inc. (1985), SAS User's Guide: Statistics, Version 5 Edition. Cary. SAS Institute Inc.

Scheuch, E., L. Graf u. S. Kühnel (1989), Volkszählung, Volkszählungsprotest und Bürgerverhalten. Ergebnisse der Begleituntersuchung zur Volkszählung 1987. Stuttgart: Metzler-Poeschel.

SPSS Inc. (1990), SPSS Reference Guide. Chicago: SPSS Inc.

Urban, D. (1990), "Multinomiale LOGIT-Modelle zur Bestimmung der Abhängigkeitsstruktur qualitativer Variablen mit mehr als zwei Ausprägungen. ZA-Information 26: 36-61.

ANHANG 1: Der Aufruf des SPSS^x-Makro "MLR"

Das für diesen Beitrag entwickelte Matrixprogramm zur ML-Schätzung multinomialer logistischer Regressionsmodelle kann als SPSS^x-Makro wie eine Prozedur aufgerufen werden. Nach dem Makronamen "MLR", der wie jedes SPSS^x-Kommando auf der ersten Spalte einer Zeile beginnen muß, folgen entsprechend der üblichen Syntax von SPSS^x eine oder mehrere Spezifikationen, die jeweils durch einen Schrägstrich getrennt werden müssen. Die Spezifikationen können auf derselben Zeile oder in Folgezeilen stehen. Sie dürfen bei Fortsetzungen über mehrere Zeilen nicht in der ersten Spalte beginnen. Im Unterschied zu Standardprozeduren können die Unterbefehle nicht auf drei Zeichen abgekürzt werden.

Die vollständige Syntax des Makros sieht folgendermaßen aus:

```
MLR  VAR = variablenliste
     /RKAT = ymax
     /NOCONST
     /COR
     /DES
     /NIT = 25     /EPS1 = 0.0000001     /EPS2 = 0.0001
     /NCASE = 20000     /PRESORT
```

Notwendig ist nach dem Makroaufruf einzig und allein die Angabe einer Variablenliste nach dem Schlüsselwort "VAR=". Für die Variablenliste gelten die üblichen SPSS^x-Konventionen. Variablennamen können durch Leerstelle oder Komma getrennt werden. Das Schlüsselwort "TO" kann verwendet werden. Gegenüber anderen Prozeduren besteht die Einschränkung, daß ein Variablenname nur einmal in der Variablenliste aufgeführt werden darf. Anderenfalls bricht das Programm mit einer SPSS^x-MATRIX-Fehlermeldung ab.

Das Matrixprogramm verlangt, daß die erste der aufgeführten Variablen die nominalskalierte abhängige Variable ist und alle folgenden Variablen die erklärenden Variablen sind. Zulässige Werte für die abhängige Variable sind nur ganze Zahlen, wobei der kleinste Wert "1" sein muß. Das Programm verlangt weiter, daß alle Kategorien von der kleinsten bis zur größten vorgefundenen ganzen Zahl Besetzungszahlen größer Null aufweisen. Gegebenenfalls muß die abhängige Variable vor dem Aufruf der Prozedur rekodiert werden. Das Makro prüft diese Bedingung ab und gibt gegebenenfalls eine entsprechende Fehlermeldung aus. Da grundsätzlich jeder Fall von der Analyse ausgeschlossen wird, wenn er auf mindestens einer der in der Variablenliste aufgeführten Variablen einen ungültigen Wert aufweist, kann ein Fehlerabbruch auch dann erfolgen, wenn eine univariate Häufigkeitsauszählung der abhängigen Variable in jeder Kategorie Besetzungen zeigt.

Das Modell der logistischen Regression unterstellt, daß die erklärenden Variablen metrisch sind. Statt einer kategorialen erklärenden Variable sind daher in der Variablenliste des Makros gegebenenfalls dichotome Dummy-Variablen aufzuführen. Dabei ist darauf zu achten, daß bei m Kategorien der erklärenden Variable maximal (m-1) Dummy-Variablen aufgeführt werden. Das Makro druckt eine Warnung aus, wenn der kleinste Eigenwert der Rohproduktmomentmatrix kleiner als 10* ist. Wenn die Matrix der zweiten Ableitungen während der Iterationen singulär wird, wird der Minimierungsprozeß mit einer entsprechenden Meldung abgebrochen.

Hinter dem Schlüsselwort "/RKAT =" kann die Referenzkategorie für die Koeffizienten der abhängigen Variable spezifiziert werden. Wird keine Referenzkategorie aufgeführt, wird als Voreinstellung der größte Wert der abhängigen Variable gewählt (oben durch "ymax" symbolisiert). Die Angabe "RKAT = 0" bewirkt, daß nacheinander jede Kategorie der abhängigen Variable als Referenzkategorie betrachtet wird und so alle theoretisch möglichen Regressionskoeffizienten ausgegeben werden.

Ergänzend zu den in der Variablenliste aufgeführten erklärenden Variablen wird standardmäßig eine Regressionskonstante in die Regressionsgleichung aufgenommen. Durch Angabe des Parameters "/NOCONST" kann ein Modell ohne Regressionskonstante geschätzt werden.

Neben den Regressionskoeffizienten und einigen daraus abgeleiteten Kenngrößen werden standardmäßig Signifikanztests der einzelnen Prädiktoren berechnet. Außerdem werden Chiquadrat-Differenzentests und Pseudo-R^2-Werte berechnet. Nullmodell ist dabei zum einen das Ausgangsmodell der Iteration, bei dem alle Regressionskoeffizienten den Wert Null haben, und zum anderen ein Modell, in dem nur die Regressionskonstante enthalten ist.[1] Zur Kontrolle des Minimierungsverlaufs wird außerdem ein Protokoll der Minimierungsschritte ausgedruckt.

Zusätzliche Ausgaben können über weitere Schlüsselwörter angefordert werden. Mit der Spezifikation "/COR" können die Korrelationen der Schätzungen ausgegeben werden. Durch Angabe des Parameters "/DES" werden deskriptive Statistiken ausgegeben. Diese umfassen eine Häufigkeitstabelle der abhängigen Variable und die Mittelwerte, Standardabweichungen und Korrelationen der erklärenden Variablen.

Drei weitere Schlüsselwörter beeinflussen den Schätzalgorithmus. Durch die Angabe von "/NIT = k" wird die maximale Anzahl von Iterationen auf k festgelegt. Als Voreinstellung sind 25 Iterationen vorgesehen. Wenn der Algorithmus nicht in der vorgegebenen Maximalzahl von Iterationen konvergiert, erfolgt eine Fehlermeldung. Die Konvergenz wird über zwei Konvergenzparameter gesteuert. Konvergenz wird zum einen angenommen, wenn der relative Abfall der Minimierungsfunktion zwischen zwei Iterationen kleiner als ein Konvergenzkriterium "EPS1" ist. Der voreingestellte Wert von 10^{-7} kann durch den Parameter "/EPS1 = ..." verändert werden. Das Matrixprogramm nimmt weiterhin Konvergenz an, wenn das betragsmäßig größte Element des Vektors der ersten Ableitungen der Regressionskoeffizienten kleiner als ein kritischer Wert "EPS2" ist Die Voreinstellung für diesen Wert ist 10^{-4} Wie das Kriterium "EPS 1" kann auch dieser Wert überschrieben werden.

Zur Berechnung von d_{it} (vgl. Gleichung 7 und 9) für jede Kategorie der abhängigen Variable muß die abhängige Variable einmalig fallweise abgearbeitet werden. Bei mehr als 20 000 Fällen muß mit der Spezifikation "/NCASE = nmax" eine größere Fallzahl angegeben werden.[2] Sind die Daten nach den Werten der abhängigen Variable aufsteigend sortiert, kann das fallweise Abarbeiten entfallen. In diesem Fall kann durch Angabe der Spezifikation "/PRESORT" die Rechenzeit beschleunigt werden.

Aufgrund der Länge des Makros ist es sinnvoll, das Matrixprogramm als externe Datei zu speichern und mit dem SPSS^x-Befehl INCLUDE bereitzustellen. Da am Anfang des Programms ein SET-Befehl ausgeführt wird, sollte vor dem INCLUDE-Befehl mit dem Befehl PRESERVE die ursprüngliche SPSS^x-Umgebung gesichert werden und nach dem Aufruf des Makros mit dem Befehl RESTORE wieder der Ursprungszustand hergestellt werden. Um den Arbeitsspeicher voll nutzen zu können, ist vor dem INCLUDE-Befehl der Befehl EXEC nützlich.

Insgesamt ergibt sich somit folgende Befehlssequenz zur Berechnung der multinomialen logistischen Regression:

```
PRESERVE
EXEC
INCLUDE FTLE="dateiname der datei mit dem MLR-makro"
MLR VAR=...
RESTORE
```

Bei einem interaktiven Arbeiten mit SPSS^X ist hinter jedem Befehl das Zeichen für den Befehlsabschluß - in der Regel ein Punkt - zu setzen. Treten beim Programmaufruf Fehlermeldungen des MATRIX-Interpreters auf, die auf Programmzeilen hinweisen, so können diese lokalisiert und möglicherweise behoben werden, wenn direkt vor dem Aufruf des Makros der SPSS^x-Befehl "SET MPR ON" eingefügt wird.

---

1 Das Ausgangsmodell entspricht der Hypothese, daß die abhängige Variable in der Population gleichverteilt ist Im Konstantenmodell wird zwar die empirische Randverteilung der abhängigen Variable reproduziert, jedoch Unabhängigkeit von den erklärenden Variablen unterstellt.

2 Der Parameter "NCASE" setzt die Anzahl der maximalen Loops im SET-Kommando von SPSS^X hoch.

ANHANG 2: SPSS^X-Anweisungen zur Schätzung der Modellparameter

```
********** MINIMIERUNG NACH NEWTON/RAPHSON-ALGORITHMUS. ************
COMPUTE FIT=CHI(1,1)                /* STARTWERT DER MIN.-FUNKTION
COMPUTE LOG(1,2)=FIT                 /* AUSGANGSFIT
COMPUTE BETA=MAKE(K,IM1,0)          /* STARTWERTE VON BETA IN 1. ITER.
COMPUTE PROB=MAKE(N,I,(1/I))        /* STARTWERTE DER WAHRSCH.
**** BEGINN MINIMIERUNGSITERATION
PRINT /TIT "*** MINIMIERUNG VON -2 LOG-LIKELIHOOD-FUNKTION"
      /SPA NEW
COMPUTE CONV=0
LOOP ITER=1 TO NITER                    /* MINIMIERUNGSSCHLEIFE
COMPUTE LOG((ITER+1),1)=ITER            /* FESTHALTEN DER ITER.-NUMMER
COMPUTE FOLD=FIT                        /* LETZTER WERT DES FITS
*** BERECHNUNG DER 1. U. 2. ABLEITUNG ***
** A. BERECHNEN DER 1. ABLEITUNG -> DEV1
COMPUTE DEV1=T(X)*(DES(:,1:IM1)-PROB(:,1:IM1))
** MINIMIERUNG ABGESCHLOSSEN, WENN 1. ABLEITUNG NULL
COMPUTE LOG((ITER+1),4)=MMAX(ABS(DEV1))
DO IF (LOG((ITER+1),4) LE EPS2)
PRINT /TITLE="***** KONVERGENZ: 1. ABLEITUNG IST NULL *****"
      /SPACE=5
COMPUTE CONV=1
END IF
** B. BERECHNEN DER INFORMATIONSMATRIX (- 2. ABLEITUNG) -> DEV2
COMPUTE DEV2=MAKE((K*IM1),(K*IM1),0)
LOOP L1=1 TO IM1
LOOP L2=1 TO L1
LOOP L3=1 TO K
DO IF (L1=1)
COMPUTE L4E=L3
ELSE
COMPUTE L4E=K
END IF
LOOP L4=1 TO L4E
DO IF (L1=L2)                           /* WENN PROB(I)=PROB(J)
COMPUTE S=PROB(:,L1)&*(1-PROB(:,L1))    /* S = VARIANZ VON P(I)
ELSE                                    /* WENN PROB(I) <> PROB(J)
COMPUTE S=-1*PROB(:,L1)&*PROB(:,L2)     /* S = KOV<PROB(I),PROB(J)>
END IF
COMPUTE DEV2((L3+K*(L1-1)),(L4+K*(L2-1)))=      /*MULTIPLIKATION MIT*/
        MSUM(X(:,L3)&*S&*X(:,L4))               /* X(K) UND X(L) */
COMPUTE DEV2((L4+K*(L2-1)),(L3+K*(L1-1)))=
        DEV2((L3+K*(L1-1)),(L4+K*(L2-1)))
END LOOP
END LOOP
END LOOP
END LOOP
RELEASE S
*** BERECHNUNG DER NEUEN BETA-WERTE UND DER LIKELIHOODFUNKTION
** A. ABBRUCH, WENN 2. ABLEITUNG NICHT POSITIV DEFINIT
DO IF (ALL(EVAL(DEV2)) = 0)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
        /SPACE=5
PRINT /TITLE="     INFORMATIONSMATRIX NICHT POSITIV DEFINIT!"
        /SPACE=1
PRINT /TIT "+++++ PROGRAMM STOPPT MINIMIERUNG!"
COMPUTE ERRMIN=1
BREAK
```

```
** B. KORREKTUMATRIX, WENN 2. ABLEITUNG POSITIV DEFINIT
                GRAD = -INV(DEV2)*DEV1
ELSE
COMPUTE DEV1=RESHAPE(T(DEV1),(K*IM1),1)
COMPUTE GRAD=INV(DEV2)*DEV1
COMPUTE GRAD=T(RESHAPE(GRAD,IM1,K))
END IF
** C. NEUE REGRESSIONSKOEFFIZIENTEN BERECHNEN
COMPUTE BETA=BETA+GRAD
** D. -2*LN LIKELIHOOD BERECHNEN
COMPUTE BEXP={ BETA, MAKE(K,1,0) }
COMPUTE PROB=(X*BEXP)
DO IF (MMAX(PROB) GT 170)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
         /SPACE=5
PRINT /TITLE="EXP(X*BETA) NICHT DARSTELLBAR!"
         /SPACE=1
COMPUTE ERRMIN=1
BREAK
END IF
COMPUTE PROB=EXP(PROB)
COMPUTE TEMP=RSUM(PROB)*MAKE(1,I,1)
COMPUTE PROB=PROB &/TEMP
COMPUTE FIT=-2*MSUM((DES&*(LN(PROB))))
RELEASE TEMP
** E. ERGEBNISS IN ITERATIONSPROTOKOLL FESTHALTEN
COMPUTE KRIT=1-(FIT/FOLD)
COMPUTE LOG((ITER+1),2)=FIT           /* FUNKTIONSWERT FESTHALTEN */
COMPUTE LOG((ITER+1),3)=KRIT          /* KRITERIUM FESTHALTEN */
*** BEWERTUNG DER ERGEBNISSE
** A. FIT HAT SICH VERSCHLECHTERT
DO IF (KRIT < 0)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
         /SPACE=5
PRINT /TITLE="     DEVIANZ HAT SICH ERHOEHT!"
         /SPACE=1
PRINT /TIT "+++++ PROGRAMM STOPPT MINIMIERUNG!"
COMPUTE ERRMIN=1
BREAK
END IF
** B. KONVERGENZ: FITZUWACHS < EPS1
DO IF (KRIT <= EPS1)
PRINT ITER /FOR "F2.0" /TITLE="***** KONVERGENZ NACH"
      /RLAB "ITERAT.:"
      /SPACE=5
COMPUTE CONV=1
END IF
** C. NAECHSTE ITERATION, WENN FITTZUWACHS > EPS1
DO IF (CONV=1)
BREAK
END IF
END LOOP        /* ENDE MINIMIERUNGSSCHLEIFE
```

Die komplementäre Anwendung zweier Verfahren: Korrespondenzanalyse und Konfigurationsfrequenzanalyse

von Jörg Blasius und Erwin Lautsch*

1. Einleitung

Nicht nur in den Sozialwissenschaften besteht ein Streit zwischen explorativem und konfirmatorischem Vorgehen bei der quantitativen Datenanalyse. Während dem einen das explorative Vorgehen als wahllos erscheint, da Zusammenhänge lediglich beschrieben, nicht aber erklärt werden können, hält der andere die Genauigkeit der mittels konfirmatorischer Verfahren gewonnenen Ergebnisse für wenig realistisch. Während der explorativ auswertende Forscher nicht selten dem Verdacht ausgesetzt ist, "keine Theorie zu haben", wird dem konfirmatorisch auswertenden häufig ein "fitting the data" vorgeworfen.

Ein weiteres Problem in den Sozialwissenschaften ist, daß die vorhandenen Daten relativ selten metrisches Meßniveau haben. Dieses wird aber für die Anwendung bei der Mehrzahl der multivariaten Analyseverfahren vorausgesetzt. Als Ausweg für diese Diskrepanz zwischen dem Meßniveau der vorhandenen Daten und den Anforderungen der Verfahren werden häufig Metrisierungen vorgenommen bzw. nicht-metrische Daten per Definition als metrisch erklärt. Im ersten Fall werden oft Skalen konstruiert, bei denen die Ursprungsdaten lediglich als Durchschnittswerte vorkommen und daher nur noch schwer interpretierbar sind, im zweiten Fall sind dagegen Fehler aufgrund der ungenauen Analyseform vorprogrammiert.

In dem vorliegenden Artikel werden zwei multivariate Verfahren vorgestellt, bei denen es keine Voraussetzungen bezüglich des Skalenniveaus gibt. Bei den Verfahren handelt es sich um die Ende der 60er Jahre von ***Lienert*** entwickelte Konfigurationsfrequenzanalyse (KEA) ***{Lienert*** 1969, ***Krauth und Lienert*** 1973, ***Lienert*** 1988, ***Lautsch und von Weber* 1990, *von Eye*** 1990) und um die erst in jüngster Zeit viel diskutierte Korrespondenzanalyse ***(Greenacre*** 1984, ***Lebart etal.*** 1984), die in den Sozialwissenschaften mit Ausnahme von französischsprachigen Publikationen bislang selten verwendet wurde (s. **z.B. *Bourdieu*** 1982, 1988, *Blasius* 1987, ***Blasius und Winkler*** 1989, ***Dangschat und Blasius*** 1987, ***Thiessen und Rohlinger*** 1988).

Während die Konfigurationsfrequenzanalyse in die Gruppe der konfirmatorischen Verfahren einzuordnen ist, zählt die Korrespondenzanalyse zu den explorativen Techniken. In diesem Beitrag sollen beide Verfahren kurz vorgestellt und Schnittpunkte der komplementären Verwendung aufgezeigt werden.

---

*Dr. Erwin Lautsch ist Priv. Doz. am Institut für Rechtswissenschaften, Otto-Nuschke-Str. 10/11, 1086 Berlin.

2. Daten

Datengrundlage ist eine schriftliche Befragung von 567 Alkoholikern des Hamburger Universitätskrankenhauses Eppendorf (UKE), die sich zwischen 1980 und 1985 beraten ließen und sich über eine Therapie informierten. Mittels dieser Befragung sollte herausgefunden werden, ob es eine bestimmte Alkoholikerpopulation gibt, deren Mitglieder überdurchschnittlich oft an einer Therapie teilnahmen und durch welche Indikatoren diese charakterisiert werden kann. Ferner sollten jene Personen beschrieben werden, die übermäßig oft nicht zur Therapie erschienen bzw. die lediglich an einer Vorbereitungsgruppe teilnahmen.

Als beschreibende Merkmale der drei genannten Personengruppen wurden verwendet: das Eingeständnis von Alkoholproblemen ("Werde damit nicht alleine fertig": stimmt; stimmt etwas; stimmt nicht), das Drängen des Partners/der Familie und der Rat des Arztes zur Therapieteilnahme (beide Variablen ebenfalls in trichotomisierter Form), der geplante Alkoholgenuß nach einer Therapie (nie wieder Alkohol; Jahre kein Alkohol; moderates Trinken), Angaben über Magenerkrankungen (nein, keine Magenerkrankung; ja, ohne Krankenhausaufenthalt; ja, mit Krankenhausaufenthalt), Kontakt zu Selbsthilfegruppen (nein; gelegentlich; regelmäßig), der Erwerbsstatus (berufstätig; arbeitslos; Hausfrau/Rentner) und der Familienstand (ledig; verheiratet; getrennt-lebend; geschieden; verwitwet).

In dem genannten Zeitraum nahmen 135 der 567 Personen des Erstgespräches (23.8%) an einer Therapie teil. Von den verbleibenden nahmen 187 (33.0%) lediglich an der Vorbereitungsgruppe teil, die anderen 245 Personen (43.2%) kamen nicht wieder.

3. Korrespondenzanalyse

Die Korrespondenzanalyse wurde unter verschiedenen Bezeichnungen in der statistischen Literatur eingeführt. In der bekanntesten und hier vorzustellenden Form wurde sie in Frankreich entwickelt (***Benzécri et al.*** 1973), sie gilt dort als sozialwissenschaftliche Standardmethode. Seit 1976 gibt es mit "Les Cahiers de l'Analyse des Données" eine Zeitschrift, die nahezu ausschließlich Artikel über die Korrespondenzanalyse (incl. zahlreicher Anwendungsbeispiele) enthält. Trotz relativ früher Schriften von ***Benzécri*** und anderen französischen Statistikern sowie zahlreicher sozialwissenschaftlicher Anwendungen, insbesondere von französischen Sozial Wissenschaftlern, wurde die Korrespondenzanalyse außerhalb von Frankreich erst wahrgenommen, als englischsprachige Literatur (insbesondere die Lehrbücher von ***Greenacre*** (1984) und ***Lebart et al*** (1984)) zur Verfügung stand. In den nächsten Jahren dürfte die Korrespondenzanalyse aufgrund der visuellen Anschaulichkeit ihrer grafischen Ausgabe und nicht zuletzt auch durch die Implementation entsprechender Prozeduren

in die neuesten Versionen der großen Statistikpakete (SAS, SPSS, BMDP) auch außerhalb von Frankreich zunehmend Anwendung finden.

Bei der Korrespondenzanalyse wird in "multiple" und "einfache" unterschieden, obwohl beide Verfahren letztlich auf dem gleichen Algorithmus basieren (vgl. *Greenacre* 1984, S. 126 ff.). Bei der einfachen Korrespondenzanalyse werden - mit einer für alle Tabellen identischen zu beschreibenden Variable - die einzelnen Kreuztabellen untereinander geschrieben. Durch diese Art der Dateneingabe gibt es keine Probleme mit der Häufigkeit der Zellenbesetzung, wie dies z.B. bei der log-linearen Analyse oder der KFA der Fall ist. Die Eingabe bei der multiplen Korrespondenzanalyse kann sowohl in Form einer Burt-Matrix erfolgen, also in Form einer symmetrischen Matrix, bei der jede Variable mit jeder anderen kreuztabelliert und die Tabellen spalten- und zeilenweise miteinander verknüpft werden, als auch in Form einer Indikatormatrix. Letztere Eingabeform wird insbesondere dann verwendet, wenn einzelne Analyseeinheiten, in der Regel Individuen, in einem mehrdimensionalen Raum lokalisiert und beschrieben werden sollen, so z.B von *Bourdieu* (1988) zur Beschreibung von Hochschullehrern im "Raum der Fakultäten". Die Verwendung von Burt-Matrices ist relativ selten; aufgrund der Duplizität von Zeilen- und Spaltenausprägungen gibt es weder eine (mehrere) zu beschreibende Variable(n) noch Individuen (oder andere Analyseeinheiten), die anhand von Merkmalen beschrieben werden sollen.

Die Korrespondenzanalyse ist ein explorativ es Analyseverfahren, statistische Signifikanztests wie bei der log-linearen Analyse oder der KFA gibt es nicht. Die Interpretation der Ergebnisse erfolgt deskriptiv - von daher sind den Einsatzmöglichkeiten der Korrespondenzanalyse Grenzen gesetzt. Sinnvoll ist ihr Einsatz aber immer dann, wenn Strukturen in mehrdimensionalen Daten gefunden werden sollen, also als hypothesengenerierendes Verfahren. Ferner kann die Korrespondenzanalyse zur vertiefenden Beschreibung von inferenzstatistisch gewonnenen Erkenntnissen verwendet werden (vgl. z.B. *van der Heijden et al.* 1989).

Die Korrespondenzanalyse ist primär eine Technik zur grafischen Darstellung von Spalten und Zeilen einer (mehrerer) Kontingenztabelle(n). Hierbei gibt es zwar eine eindeutige Distanzfunktion zwischen den Zeilen und zwischen den Spalten, nicht aber zwischen Spalten **und** Zeilen (vgl. *Greenacre und Hastie* 1987, *Greenacre* 1989). Wie bei der Hauptkomponentenanalyse gibt es einen Satz von orthogonalen Vektoren, die einen niederdimensionalen Raum (maximal (r-1) mit $r \leq c$, wobei r=Anzahl der Spalten und c=Anzahl der Zeilen) aufspannen. Neben der grafischen Darstellung der Ergebnisse gibt es eine numerische Ausgabe, mittels derer eine weitergehende Interpretation der Daten möglich ist. Mit Hilfe der dazugehörigen Koeffizienten ist es möglich, Zeilen und Spalten der zu analysierenden Kontingenztabelle(n) zu vergleichen.

4. Konfigurationsfrequenzanalyse (KFA)

Die KFA wurde von ***Lienert*** (1969) ursprünglich als Klassifikationsmethode entwickelt. Sie ist eine unter mehreren Zugängen, die zu einer bestimmten Typologie von Individuen führt, wobei die Besonderheit der KFA in der inferenzstatistischen Anwendungsmöglichkeit besteht.

Im klassischen Modell der KFA sollen Individuen mit identischen Merkmalsausprägungen (dargestellt durch die einzelnen Zellen einer mehrdimensionalen Kontingenztabelle) als "typische Klasse" mittels einer statistischen Testentscheidung identifiziert werden. So kann z.B. geprüft werden, ob Personen mit dem Merkmalsprofil "geplanter zukünftiger Alkoholgenuß: Abstinenz", "der Partner/die Familie drängt stark zur Therapie" und "Magenleiden ja, aber ohne Krankenhausaufenthalt" eine statistisch bedeutsame Klasse manifestieren. In dem genannten Beispiel wäre eine dreidimensionale Kontingenztabelle zu konstruieren, deren Zellen aus allen möglichen Merkmalskombinationen der drei Variablen (also 27) bestehen.

Entsprechend dem Modell der KFA ist eine Personengruppe mit identischen Merkmalsausprägungen "statistisch bedeutsam", wenn unter der Annahme der statistischen Unabhängigkeit der Variablen die beobachtete Zellenhäufigkeit statistisch signifikant von der erwarteten abweicht. Für die konkrete Testentscheidung gibt es mehrere Testprozeduren, von denen in den meisten Fällen die nach Lehmacher (mit Stetigkeitskorrektur und Alpha-Adjustierung) zu empfehlen ist ***(Lautsch und von Weber*** 1990). Merkmalskombinationen, die mit dem Modell der statistischen Unabhängigkeit nicht vereinbar sind, werden als "Typen" bzw. als "Antitypen" bezeichnet. Wir sprechen von einem "Typ", wenn die beobachtete Zellenhäufigkeit größer als die der erwarteten ist, und von einem "Antityp", wenn die beobachtete Häufigkeit kleiner als die erwartete ist.

Der Unterschied zur log-linearen Analyse besteht darin, daß die KFA ausschließlich zellenorientiert ist, d.h. es werden Merkmalskombinationen gesucht, die unter den o.g. Annahmen als Ausreißer interpretiert werden können. Eine Zerlegung der Varianz in Haupteffekte und Interaktionseffekte unterschiedlicher Ordnung sowie die Prüfung bestimmter Modellannahmen sind bei der standardmäßigen Anwendung der KFA nicht vorgesehen. Mittels Weiterentwicklungen der KFA ist es inzwischen möglich, auch spezifische Klassifikationsprobleme zu lösen (z.B. durch die hierarchische KFA, die Mehrstichproben-KFA, die Prädiktions-KFA). In empirischen Anwendungen werden diese Verfahren in der Regel als nicht-parametrische Pendants klassischer multivariater Verfahren verwendet (vgl. ***Lienert*** 1988, *Lautsch* ***und von Weber*** 1990, ***von Eye*** 1990).

Um zu erklären, welche der drei Therapiekategorien mit welchen beschreibenden Merkmalskombinationen einhergehen, verwenden wir ein modifiziertes Modell - die (multiple) Prädiktions-KFA (vgl. ***Lienert*** 1988). Die hierfür zu konstruierende Kontingenztabelle hat zwei Dimensionen: Die erste ist durch die drei Merkmalsausprägungen der abhängigen Größe (Therapieteilnahme) gegeben, die zweite durch die Merkmalsausprägungskombinationen der unabhängigen (der beschreibenden) Variablen. Ein "Prädiktionstyp" bzw. ein "Prädiktionsantityp" liegt vor, wenn eine Kombination von Merkmalsausprägungen der unabhängigen Variablen mit einer Ausprägung der abhängigen Variable signifikant häufiger bzw. signifikant seltener als erwartet (unter der Annahme der statistischen Unabhängigkeit) vorkommt. Während beim KFA-Basismodell alle untersuchten Merkmale gleichberechtigt in die Analyse eingehen (symmetrisches Modell), wird bei der Prädiktions-KFA nach zu erklärenden (abhängigen) und erklärenden (unabhängigen) Merkmalen (asymmetrisches Modell) unterschieden. Zellenweise getestet wird bei der Prädiktions-KFA die Nullhypothese, daß zwischen dem zu erklärenden Merkmal "Therapieteilnahme" und den erklärenden Merkmalen (den Merkmalskombinationen derselbigen) kein Zusammenhang besteht.

Mit steigender Anzahl der beschreibenden Variablen steigt die Anzahl der Kategorien der zu bildenden Merkmalskombinationen (die resultierende Variable wird auch als "synthetische Variable" bezeichnet) schnell an. So gibt es bei 5 Merkmalen mit je 3 Ausprägungen 243 mögliche Merkmalskombinationen - eine (Prädiktions-)KFA wäre damit nur noch bei einer sehr hohen Fallzahl durchführbar. Aus diesem Grund muß die Anzahl der beschreibenden Merkmale auf jene Variablen reduziert werden, die wesentlich zur Erklärung des abhängigen Merkmales beitragen. Ein mögliches Hilfsmittel bei der Suche der "wesentlichen Variablen" ist die Korrespondenzanalyse.

5. Verbindung von KFA und Korrespondenzanalyse

Soll einerseits nicht auf die Anschaulichkeit der Korrespondenzanalyse verzichtet werden, und soll andererseits eine Interpretation im inferenzstatistischen Sinn möglich sein, so können beide Verfahren ergänzend verwendet werden. Ein analoges Vorgehen wählten u.a. ***van der Heijden und de Leeuw*** (1985), ***Goodman*** (1986) und ***van der Heijden et al.*** (1989) bei ihrer komplementären Verwendung von log-linearen Modellen und der Korrespondenzanalyse.

Bei einer kombinierten Anwendung eines explorativen und eines konfirmatorischen Verfahrens kann - sofern die (wenigen) relevanten Variablen bekannt sind - im ersten Schritt das inferenzstatistische Verfahren (entsprechend der Fragestellung z.B. die log-lineare Analyse) zur Verifizierung von Interaktionen zwischen den Variablen verwendet werden. Im zweiten

Schritt kann dann die Korrespondenzanalyse auf zweifache Art eingesetzt werden: Zum einen zur explorativen Beschreibung von Variablenzusammenhängen, die zwar theoretisch relevant sind, aber im Test nicht fitten, und zum anderen zur Interpretation von Interaktionsparametern, die nur schwach signifikant sind.

Eine weitere Möglichkeit der gemeinsamen Verwendung der Verfahren besteht darin, im ersten Schritt mittels der Korrespondenzanalyse nicht-triviale Strukturen in mehrdimensionalen Daten zu explorieren. Im zweiten Schritt können die gefundenen Zusammenhänge (anhand eines weiteren Datensatzes) auf Signifikanz getestet werden.

Im folgenden wird zunächst mittels der Korrespondenzanalyse die zu beschreibende Variable "Therapieteilnahme" mit den drei Ausprägungen "nur am Erstgespräch teilgenommen", "nur an der Vorbereitungsgruppe teilgenommen" und "Therapieteilnehmer" mit Hilfe der im Abschnitt "Daten" genannten Variablen beschrieben. Daran anschließend werden die explorativ gefundenen bedeutsamen unabhängigen Variablen mit der Variable "Therapieteilnahme" kreuztabelliert und geprüft, ob es "Prädiktionstypen" oder "Prädiktionsantitypen" im Sinne der Prädiktions-KFA gibt. Im letzten Schritt werden die nachgewiesenen "Prädiktionstypen" und "Prädiktionsantitypen" zur besseren Anschaulichkeit, und zur Darstellung der (Un-)Ähnlichkeiten zwischen ihnen, mittels der Korrespondenzanalyse in einem zweidimensionalen Unterraum abgebildet.

6. Erster Schritt: Ergebnisse der ersten Korrespondenzanalyse

Eingangsmaterial der Korrespondenzanalyse sind die zeilenweise untereinandergeschriebenen absoluten Häufigkeiten der Kreuztabellen der Variablen "Therapieteilnahme" mit den im Abschnitt 2 aufgeführten Variablen.

Tabelle 1: Eingangsdaten der ersten Korrespondenzanalyse

	nur Erst-gespräch	nur Vb.-gruppe	Therapie-teilnähme
Berufstätig	108	84	58
Arbeitslos	110	75	61
Hausfrau/Rentner	26	26	15
Keine Alkoholprobleme	37	19	3
Etwas Alkoholprobleme	53	36	15
Starke Alkoholprobleme	130	122	116
Familie drängt nicht	97	57	55
Familie drängt etwas	49	48	29
Familie drängt stark	70	66	42
Arzt drängt nicht	61	38	24
Arzt drängt etwas	46	30	12
Arzt drängt stark	117	106	95
Abstinenz	158	132	109
Jahre ohne Alkohol	25	19	14
Moderates Trinken	55	31	11
Kein Magenleiden	152	114	67
Magenleiden ohne K.haus	27	26	41
Magenleiden mit K.haus	26	24	11
Selbsthilfegruppe nie	128	83	56
Selbsthilfegruppe geleg.	78	68	44
Selbsthilfegruppe regel.	36	25	32
Ledig	73	50	34
Verheiratet	88	63	46
Getrennt-lebend	22	25	11
Geschieden	56	47	40
Verwitwet	6	2	4

Für eine erste Bewertung der in den Daten vorhandenen Varianz kann die Chi-Quadrat-Statistik herangezogen werden. In Tabelle 2 sind die einzelnen quadrierten Abweichungen der empirischen von den erwarteten Werten, dividiert durch die jeweiligen Erwartungswerte, aufgeführt. Die Summe der in Tabelle 2 aufgeführten Werte ist der Gesamt-Chi-Quadrat-Wert

(Chi^2=110.5 mit df=50). Dieser Wert wird durch das Gesamt-N (das N von Tabelle 1) dividiert (der resultierende Wert wird als "total inertia" bezeichnet; total inertia=0.0257), der resultierende Wert entspricht der Spur der zu zerlegenden Residualmatrix und damit der Summe der Eigenwerte, die zu den den Projektionsraum aufspannenden Eigenvektoren gehört (vgl. ***Blasius und Rohlinger*** 1989, S. 100ff.).

Tabelle 2: Chi-Quadrat-Statistik

	nur Erst-gespräch	nur Vb.-gruppe	Therapie-teilnähme
Berufstätig	0.01	0.03	0.13
Arbeitslos	0.23	0.46	0.02
Hausfrau/Rentner	0.24	0.69	0.10
Keine Alkoholprobleme	5.53	0.01	8.98
Etwas Alkoholprobleme	1.66	0.09	4.20
Starke Alkoholprobleme	4.69	0.00	7.82
Familie drängt nicht	0.67	2.06	0.34
Familie drängt etwas	0.43	1.00	0.09
Familie drängt stark	0.47	0.91	0.04
Arzt drängt nicht	1.37	0.16	1.17
Arzt drängt etwas	1.89	0.03	4.14
Arzt drängt stark	2.60	0.01	4.02
Abstinenz	0.90	0.00	1.46
Jahre ohne Alkohol	0.00	0.00	0.00
Moderates Trinken	4.45	0.03	6.73
Kein Magenleiden	0.68	0.16	2.43
Magenleiden ohne K.haus	4.30	0.80	14.37
Magenleiden mit Khaus	0.00	0.75	0.99
Selbsthilfegruppe nie	1.72	0.29	1.24
Selbsthilfegruppe geleg.	0.12	0.46	0.11
Selbsthilfegruppe regel.	0.35	1.05	3.88
Ledig	0.53	0.06	0.46
Verheiratet	0.18	0.06	0.08
Getrennt-lebend	0.31	1.81	0.69
Geschieden	0.42	0.00	0.78
Verwitwet	0.15	0.97	0.40

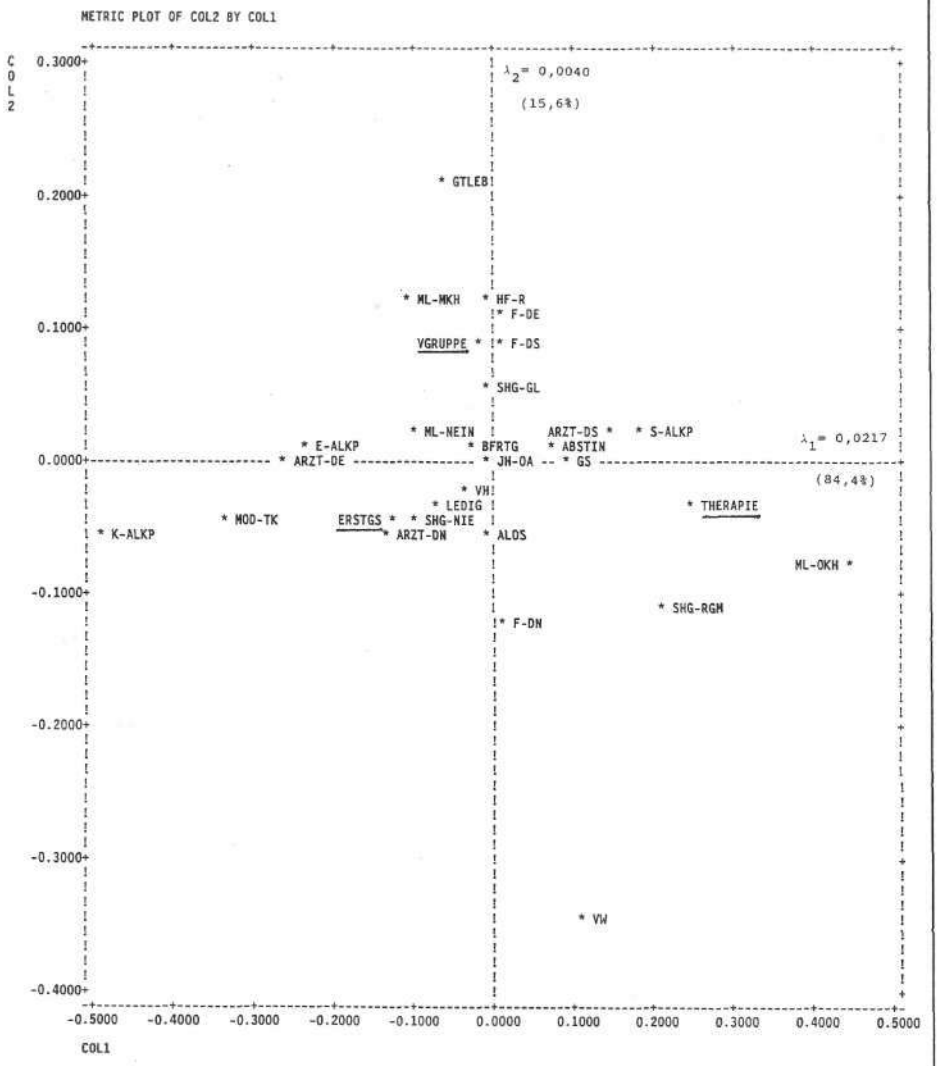
Anhand von Tabelle 2 wird ersichtlich, daß die größten Abweichungen von empirischen und erwarteten Werten in der Spalte "Therapieteilnehmer" mit den Variablenausprägungen "Magenleiden, aber ohne Krankenhausaufenthalt", "keine Alkoholprobleme" und "starke Alkoholprobleme" sind. Die Zellen dieser drei Variablenkombinationen weisen somit die größte Varianz auf. Dementgegen hat die Ausprägung "Jahre ohne Alkohol" keine Varianz in der zu erklärenden Variable - die unter der Annahme der statistischen Unabhängigkeit berechneten Erwartungswerte sind gleich denen der beobachteten. Schon an dieser Stelle wird deutlich, welche Merkmalskombinationen miteinander korrelieren (positiv, wenn der Erwartungs wert kleiner dem empirischen ist oder negativ, wenn der Erwartungswert größer als der empirische ist).

Neben den bereits erwähnten Korrelationen des Merkmales "Therapieteilnehmer" dürfte es Zusammenhänge zwischen dem zu erklärenden Merkmal "nur an der Vorbereitungsgruppe teilgenommen" und den beschreibenden Merkmalen "Familie drängt nicht", "Familie drängt etwas", "Familie drängt stark", "getrennt-lebend", "gelegentliche Teilnahme an einer Selbsthilfegruppe" und "verwitwet" geben. Im Vergleich zu den Abweichungen in der Spalte "Therapieteilnehmer" sind diese jedoch wesentlich niedriger, d.h. der Anteil der zu erklärenden Gesamtvarianz ist wesentlich geringer. In der noch durchzuführenden Korrespondenzanalyse wird sich dies dahingehend widerspiegeln, daß die Ausprägung "nur an der Vorbereitungsgruppe teilgenommen" gegenüber der Ausprägung "Therapieteilnehmer" genau den Anteil zur Determination des Darstellungsraumes beiträgt, der dem Verhältnis der beiden spaltenweise auf summierten Abweichungen entspricht (64.67 zu 11.89).

Abbildung 1: Legende

ERSTGS	nur Erstgespräch	ARZT-DS	Arzt drängt stark
VGRUPPE	nur Vorbereitungsgruppe	ABSTIN	Abstinenz
THERAPIE	Therapieteilnehmer	JH-OA	Jahre ohne Alkohol
		MOD-TK	Moderates Trinken
BFRTG	Berufstätig	ML-NEIN	Kein Magenleiden
ALOS	Arbeitslos	ML-OKH	Magenleiden ohne Knaus
HF-R	Hausfrau/Rentner	ML-MKH	Magenleiden mit Knaus
K-ALKP	Keine Alkoholprobleme	SHG-NJE	Selbsthilfegruppe nie
E-ALKP	Etwas Alkoholprobleme	SHG-GL	Selbsthilfegruppe geleg.
S-ALKP	Starke Alkoholprobleme	SHG-RGM	Selbsthilfegruppe regel.
F-DN	Familie drängt nicht	LEDIG	Ledig
F-DE	Familie drängt etwas	VH	Verheiratet
F-DS	Familie drängt stark	GTLEB	Getrennt-lebend
ARZT-DN	Arzt drängt nicht	GS	Geschieden
ARZT-DE	Arzt drängt etwas	VW	Verwitwet

Abbildung 1: Grafische Darstellung der ersten Korrespondenzanalyse*



*Zur besseren Übersicht wurden die Ausprägungen der zu erklärenden Variablen "Therapieteilnahme" unterstrichen.

Wie aus Abbildung 1 ersichtlich wird, bilden auf der Ebene der Variablen "Therapieteilnehmer" die Teilnehmer (THERAPIE) und die Nichtteilnehmer (ERSTGS) die gegensätzlichen Positionen auf der ersten Achse, diese erklärt 84.4% der Gesamtvariation des Modells. Die zweite Achse erklärt die verbleibenden 15.6% der Gesamtvariation des Modells. Sie ist auf der Ebene der Spaltenvariablen durch die Ausprägung "nur zur Vorbereitungsgruppe erschienen" gekennzeichnet.

Dem positiven Abschnitt der ersten Achse, also dem Abschnitt, dem auch die Ausprägung "Therapieteilnehmer" zuzuordnen ist, lassen sich die beschreibenden Merkmale "Magenleiden ja, aber ohne Krankenhausaufenthalt" (ML-OKH), "geschieden" (GS) und "zukünftiges Trinkverhalten: Abstinenz" (ABSTIN) zuordnen. Im gleichen Achsenabschnitt lassen sich dem zu beschreibenden Merkmal "nur am Erstgespräch teilgenommen" folgende beschreibende Merkmale zurechnen: Arzt drängt etwas (ARZT-DE), etwas Alkoholprobleme (E-ALKP), keine Teilnahme an Selbsthilfegruppen (SHG-NIE), die Familienstände "ledig" (LEDIG) und "verheiratet" (VH), zukünftig moderates Trinkverhalten (MOD-TK) und keine Alkoholprobleme (K-ALKP). Dem positiven Abschnitt der zweiten Achse, also dem gleichen Abschnitt wie dem Merkmal "nur an der Vorbereitungsgruppe teilgenommen", lassen sich die Variablenausprägungen "Hausfrau/Rentner" (HF-R), "gelegentliche Teilnahme an Selbsthilfegruppen" (SHG-GL) sowie die beiden Merkmale "Familie/Partner drängt etwas" (F-DE) und "Familie/Partner drängt stark" (F-DS) zuordnen.

Wie bereits erwähnt, wurde bislang noch keine eindeutige Distanzfunktion gefunden, mittels derer es möglich ist, Zeilen- und Spaltenausprägungen direkt miteinander zu vergleichen. Somit sollte die grafische Darstellung vorsichtig interpretiert werden. Um (genauere) Aussagen über die (Un-)Ähnlichkeit von Spalten- und Zeilenausprägungen machen zu können, kann ein Vergleich über die Ähnlichkeit der Winkel zum Schwerpunkt (dargestellt durch das Achsenkreuz) erfolgen. Diese Winkel zum Achsenkreuz, oder genauer deren quadrierte Kosinusse, sind neben anderen Koeffizienten in der numerischen Ausgabe der Korrespondenzanalyse enthalten (s. Tabelle 3). In den Spalten 5 und 8 (QCOR1, QCOR2) stehen die quadrierten Korrelationen (die quadrierten Kosinusse der Winkel zwischen den Eigenvektoren und den vom Schwerpunkt ausgehenden Vektoren der Merkmale) der Variablenausprägungen von Zeilen und Spalten mit den beiden Achsen. Die Zeilensummen der quadrierten Korrelationen der berücksichtigten Achsen sind in der zweiten Spalte (SQCOR) aufgeführt. Die Komponenten von SQCOR sind analog den Kommunalitäten der Hauptkomponentenanalyse interpretierbar, die von QCOR1 und QCOR2 denen der dazugehörigen (quadrierten) Faktorladungen. Da in unserem Beispiel nur zwei Dimensionen existieren, ergeben die in SQCOR aufgeführten Zeilensummen immer Eins. Die Spalten QCOR1 und QCOR2 sind die wichtigsten Interpretationshilfen, da hier abzulesen ist, auf welcher Achse welche Variablenausprägungen positiv oder negativ "laden" (das Vorzeichen ist den Spalten LOCI und LOC2 zu entnehmen).

Tabelle 3: Numerische Ausgabe der ersten Korrespondenzanalyse

GENSTAT	MASS	SQCOR	INR	LOCI	QCOR1	INR1	LOC2	QCOR2	INR2
ERSTGS	0.427	1.000	0.307	-0.127	0.870	0.316	-0.049	0.130	0.257
VGRUPPE	0.330	1.000	0.108	-0.017	0.036	0.005	0.090	0.964	0.666
THERAPIE	0.243	1.000	0.585	0.246	0.979	0.679	-0.036	0.021	0.078
BFRTG	0.058	1.000	0.002	-0.024	0.813	0.002	0.011	0.187	0.002
ALOS	0.057	1.000	0.006	-0.007	0.015	0.000	-0.054	0.985	0.041
HF-R	0.016	1.000	0.009	-0.006	0.002	0.000	0.124	0.998	0.060
K-ALKP	0.014	1.000	0.131	-0.493	0.987	0.154	-0.057	0.013	0.011
E-ALKP	0.024	1.000	0.054	-0.239	0.996	0.063	0.015	0.004	0.001
S-ALKP	0.086	1.000	0.113	0.183	0.989	0.133	0.019	0.011	0.008
F-DN	0.049	1.000	0.028	0.008	0.004	0.000	-0.121	0.996	0.177
F-DE	0.029	1.000	0.014	0.005	0.002	0.000	0.110	0.998	0.088
F-DS	0.041	1.000	0.013	0.012	0.018	0.000	0.089	0.982	0.081
ARZT-DN	0.029	1.000	0.024	-0.137	0.859	0.025	-0.056	0.141	0.022
ARZT-DE	0.020	1.000	0.055	-0.262	1.000	0.065	0.002	0.000	0.000
ARZT-DS	0.074	1.000	0.060	0.143	0.981	0.070	0.020	0.019	0.007
ABSTIN	0.093	1.000	0.021	0.076	0.987	0.025	0.009	0.013	0.002
JH-OA	0.014	1.000	0.000	-0.006	0.622	0.000	-0.005	0.378	0.000
MOD-TK	0.023	1.000	0.101	-0.336	0.979	0.118	-0.050	0.021	0.014
ML-NEIN	0.078	1.000	0.030	-0.097	0.964	0.034	0.019	0.036	0.007
ML-OKH	0.022	1.000	0.176	0.449	0.972	0.203	-0.076	0.028	0.031
ML-MKH	0.014	1.000	0.016	-0.112	0.439	0.008	0.127	0.561	0.057
SHG-NIE	0.062	1.000	0.029	-0.099	0.807	0.028	-0.048	0.193	0.036
SHG-GL	0.044	1.000	0.006	-0.009	0.021	0.000	0.059	0.979	0.039
SHG-RGM	0.022	1.000	0.048	0.210	0.777	0.044	-0.112	0.223	0.068
LEDIG	0.037	1.000	0.010	-0.076	0.863	0.010	-0.030	0.137	0.008
VH	0.046	1.000	0.003	-0.032	0.646	0.002	-0.024	0.354	0.006
GTLEB	0.014	1.000	0.025	-0.061	0.076	0.002	0.211	0.924	0.150
GS	0.033	1.000	0.011	0.091	0.996	0.013	0.005	0.004	0.000
VW	0.003	1.000	0.014	0.107	0.090	0.001	-0.339	0.910	0.080

In der ersten Spalte (MASS) ist die relative Masse der einzelnen Variablenausprägungen angegeben, deren Werte für Zeilen und Spalten getrennt auf "Eins" normiert wurden. Hier ist z.B. abzulesen, daß die "Therapieteilnehmer" die Masse .243 haben, d.h. ihr Anteil an allen Nennungen der berücksichtigten Variablen beträgt 24.3%. Zusätzlich zu den analogen Koeffizienten der Hauptkomponentenanalyse gibt es bei der Korrespondenzanalyse das "Trägheitsgewicht" (INR1, INR2). Dieses wird berechnet aus dem Quadrat der Entfernung der Projektion auf der Achse zum Achsenkreuz, multipliziert mit der Masse. Anhand der Spalten INR1 und INR2 wird ersichtlich, wie stark die einzelnen Achsen von den einzelnen Variablenausprägungen, getrennt für Zeilen und Spalten auf Eins normiert, determiniert werden. Die gleiche Interpretation, bezogen auf das Gesamtmodell, gilt auch für das Gesamtträgheitsgewicht (INR). Die Komponenten dieser Spalte ergeben sich aus den mit den erklärten Varianzen der Eigenwerte multiplizierten "Trägheitsgewichten" der einzelnen Achsen. Die Werte dieser Spalte entsprechen auch dem (normierten) Verhältnis der zeilen- bzw. spaltenweise aufsummierten Werte der Chi-Quadrat-Statistik (vgl. Tabelle 2).

Aus Tabelle 3 wird ersichtlich, daß - wie schon anhand der grafischen Darstellung zu erkennen war - mit der ersten Achse die Variablenausprägungen "Therapieteilnehmer" (QCORl=.979) und "nur am Erstgespräch teilgenommen" (QCOR1=.870) mit unterschiedlichen Vorzeichen (vgl. Spalte LOCI) korrelieren; d.h. beide Merkmale korrelieren hinsichtlich der verwendeten beschreibenden Variablen negativ. Ferner sind sie, entsprechend dem Modell der Korrespondenzanalyse, mit jenen Merkmalen unkorreliert, die auf einer anderen Achse laden.

Aufgespannt wird der Raum auf der Ebene der zu beschreibenden Variable insbesondere durch die Ausprägung "Therapieteilnehmer" auf der ersten Achse (INR1=.679) - diese Variablenausprägung erklärt auf der Ebene der Spalten 67.9% der Gesamtvariation der ersten Achse sowie durch die Variablenausprägung "nur an der Vorbereitungsgruppe teilgenommen" auf der zweiten Achse (INR2=.666). Auf der Ebene der beschreibenden Variablen wird die geometrische Ausrichtung der ersten Achse durch die Ausprägungen "habe keine Alkoholprobleme" (INR 1=. 154), "habe starke Alkoholprobleme" (INR1=. 133), "zukünftiges Trinkverhalten: moderat" (INR 1=. 118) und "Magenleiden ja, aber ohne Krankenhausaufenthalt" (INR 1 =.203) determiniert. Die zweite Achse wird auf der Ebene der beschreibenden Variablen insbesondere durch die drei Ausprägungen der Variablen "Familie/ Partner drängt zur Therapie" (diese drei Ausprägungen erklären zusammen 34.6% der geometrischen Ausrichtung der zweiten Achse) und durch das Merkmal "getrennt- lebend" (INR2=.150) erklärt.

Mit dem positiven Abschnitt der ersten Achse - also mit dem Abschnitt, mit dem die Gruppe der Therapieteilnehmer beschrieben werden kann - gibt es quadrierte Korrelationen von über .90 mit den Merkmalen "habe starke Alkoholprobleme", "Arzt drängt stark", "zukünftiges

Trinkverhalten: Abstinenz", "Magenleiden ja, aber ohne Krankenhausaufenthalt" und "geschieden". Mit dem negativen Abschnitt der ersten Achse, also dem Abschnitt, auf dem auch die Variablenausprägung "nur am Erstgespräch teilgenommen" lädt, gibt es quadrierte Korrelationen von über .90 mit den Variablenausprägungen "habe keine Alkoholprobleme", "habe etwas Alkoholprobleme", "Arzt drängt nicht", "Arzt drängt etwas", "zukünftiges Trinkverhalten: moderat" und "kein Magenleiden". Diese zuletzt genannten Merkmale sind mit dem Merkmal "Therapieteilnehmer" negativ korreliert, d.h. diese Merkmale treffen für diese Gruppe von Personen überdurchschnittlich selten zu.

Entsprechend dem Modell der Korrespondenzanalyse sind die Achsen unkorreliert. Zusammen mit dem beschreibenden Merkmal "nur an der Vorbereitungsgruppe teilgenommen" korrelieren folgende Variablenausprägungen mit dem positiven Abschnitt der zweiten Achse: "Hausfrau/Rentner", "Familie/Partner drängt etwas", "Familie/Partner drängt stark", "gelegentliche Teilnahme an Treffen von Selbsthilfegruppen" und "getrennt-lebend". Negativ mit den genannten Merkmalen, also mit dem negativen Abschnitt der zweiten Achse, korrelieren die Variablenausprägungen "arbeitslos", "die Familie/der Partner drängt nicht" und "verwitwet".

Diesen Ergebnissen zufolge sind akute Magenleiden, die zum Befragungszeitpunkt noch nicht im Krankenhaus behandelt wurden, der wichtigste Indikator zur Therapieteilnahme. Ebenfalls einen hohen positiven Einfluß auf die Therapieteilnahme hat die Einsicht, daß "starke Alkoholprobleme" bestehen. Wird von dem Patienten jedoch angegeben, daß "keine Alkoholprobleme bestehen" bzw. daß sein zukünftiges Trinkziel "moderates Trinken" ist, so ist die Wahrscheinlichkeit hoch, daß er nach dem Erstgespräch nicht wiederkommt. Ferner hat das "Drängen der Familie/des Partners" keinen Einfluß auf die Entscheidung für oder gegen eine Therapieteilnahme, es führt lediglich dazu, daß der Patient überdurchschnittlich häufig ausschließlich an der Vorbereitungsgruppe teilnimmt.

Die Variable, deren Ausprägungen alle mit der ersten Achse hoch korrelieren, ist "Alkoholprobleme". Zudem hat sie mit insgesamt 35.3% den von allen Variablen größten Anteil an der geometrischen Ausrichtung der ersten Achse. Analoges gilt in bezug auf die zweite Achse für die Variable "die Familie/der Partner drängt". Somit scheinen beide Variablen gut für die Beschreibung der Variable "Therapieteilnahme" geeignet zu sein.

Im folgenden Abschnitt soll die Frage beantwortet werden, ob den Therapieteilnehmern, denjenigen, die nur die Vorbereitungsgruppe besuchten und denjenigen, die nur zum Erstgespräch kamen, statistisch signifikante unterschiedliche Merkmalskombinationen der beiden ausgewählten Variablen zugewiesen werden können, die diese im Sinne der Prädiktions-KFA als "Prädiktionstypen" oder als "Prädiktionsantitypen" kennzeichnen.

7. Zweiter Schritt: Konfigurationsfrequenzanalyse

Zur Durchführung der Prädiktions-KFA wird eine zweidimensionale Kreuztabelle mit der abhängigen Merkmalsgröße "Therapieteilnahme" einerseits und der aus den Merkmalen "Eingeständnis von Alkoholproblemen" und "Familie/Partner drängt zur Therapieteilnahme" kombinierten Variablen andererseits gebildet[1].

Tabelle 4: Eingabedaten der Prädiktions-KFA

Alkohol-probleme[1]	Familie drängt[2]	nur Erst-gespräch	nur Vb.-gruppe	Therapie-teilnahme
1	1	21	5	3
1	2	9	4	0
1	3	7	10	0
2	1	18	10	5
2	2	12	9	1
2	3	17	16	9
3	1	57	41	47
3	2	27	34	28
3	3	42	40	33

[1] (1) keine Alkoholprobleme, (2) etwas Alkoholprobleme, (3) starke Alkoholprobleme
[2] (1) drängt nicht, (2) drängt etwas, (3) drängt stark

Die in Tabelle 4 aufgeführten absoluten Häufigkeiten sind das Eingangsmaterial für die Prädiktions-KFA. Für alle Zellen der Kontingenztabelle wird als Testgröße Lehmachers asymptotischer hypergeometrischer Komponententest mit der Kontinuitätskorrektur nach Küchenhoff verwendet (s. *Lautsch und von Weber* 1990, S. 60ff.). Die Testentscheidungen werden zum einen explorativ (zweiseitige Testentscheidung ohne Alpha-Adjustierung) und zum anderen inferenzstatistisch (einseitige Testentscheidung mit Alpha-Adjustierung nach Holms) durchgeführt[2] (s. *Lautsch und von Weber* 1990, S. 69f.). Der kritische Wert liegt beim zweiseitigen Test bei u=1.96 (5%-Niveau), beim einseitigen Test mit Alpha-Adjustierung nach Holms gibt es keinen konstanten Wert.

1 Damit die Unabhängigkeit der Ergebnisse, die mit den unterschiedlichen Verfahren erzielt wurden, gewährleistet ist, hätte vor Beginn der Datenauswertung die Stichprobe gesplittet werden müssen. Da wir aber lediglich eine Auswertungsstrategie angeben und zudem das Leistungsvermögen beider Verfahren diskutieren wollen, wurde auf ein Splitten der Stichprobe verzichtet

2 Als Programm zur KFA wurde SICFA verwendet *(Lautsch und von Weber* 1990).

Tabelle 5: Ergebnisse der Prädiktions-KFA

Alkohol-probleme[1]	Familie drängt[2]	Therapie-teilnahme[3]	emp. Wert	erw. Wert	Lehmacher-Test[4] ohne α-Adj.	mit α-Adj.
1	1	1	21	12.1	2.80 (1)	3.27 (1)
1	1	2	5	9.7	1.99(-1)	1.70
1	1	3	3	7.2	1.97(-1)	1.65
1	2	1	9	5.4	1.94	1.76
1	2	2	4	4.4	1.28	.00
1	2	3	0	3.2	2.04(-1)	1.78
1	3	1	7	7.1	1.42	.00
1	3	2	10	5.7	2.11 (1)	1.99
1	3	3	0	4.2	2.44(-1)	2.13
2	1	1	18	13.7	1.93	1.38
2	1	2	10	11.0	1.54	.21
2	1	3	5	8.2	1.70	1.14
2	2	1	12	9.1	1.75	1.04
2	2	2	9	7.4	1.57	.53
2	2	3	1	5.5	2.23(-1)	2.01
2	3	1	17	17.5	1.64	.00
2	3	2	16	14.1	1.68	.49
2	3	3	9	10.5	1.54	.36
3	1	1	57	60.3	1.93	.56
3	1	2	41	48.5	2.11(-1)	1.46
3	1	3	47	36.2	2.49 (1)	2.34
3	2	1	27	37.0	2.40(-1)	2.25
3	2	2	34	29.8	1.92	.92
3	2	3	28	22.2	2.06 (1)	1.43
3	3	1	42	47.8	2.00(-1)	1.14
3	3	2	40	38.5	1.83	.23
3	3	3	33	28.7	1.92	.93

[1] (1) keine Alkoholprobleme, (2) etwas Alkoholprobleme, (3) starke Alkoholprobleme

[2] (1) drängt nicht, (2) drängt etwas, (3) drängt stark

[3] (1) nur Erstgespräch, (2) nur an Vorbereitungsgruppe teilgenommen, (3) Therapieteilnahme

[4] (1) "Prädiktionstyp", (-1) "Prädiktionsantityp"

Gesamt-Chi^2=43.5 mit df=16

Aus Tabelle 5 wird ersichtlich, daß es bei inferenzstatistischen Testentscheidungen (mit Alpha-Adjustierung) nur einen "Prädiktionstyp" gibt. Dieser ist dadurch charakterisiert, daß die zusammengesetzte Variablenausprägung "keine Alkoholprobleme, die Familie/der Partner drängt nicht zur Therapieteilnahme" einen signifikanten Einfluß auf die Variablenausprägung "nur am Erstgespräch teilgenommen" hat. Weitere "Prädikationstypen" sowie auch "Prädikationsantitypen" sind lediglich tendenziell, d.h. wenn auf die Alpha-Adjustierung verzichtet wird, vorhanden. Demnach käme es relativ oft nur zur ausschließlichen Teilnahme am Erstgespräch, wenn angegeben wurde, daß keine Alkoholprobleme bestehen und daß die Familie/der Partner nur etwas zur Therapie drängte. Relativ häufig zur Teilnahme an der Vorbereitungsgruppe, aber nicht an der Therapie, käme es, wenn die Familie/der Partner zwar stark zur Therapieteilnahme drängte, aber gleichzeitig "keine Alkoholprobleme" angegeben wurde. Tendenzielle Charakteristika für eine Therapieteilnahme sind die Angaben "habe starke Alkoholprobleme, die Familie/der Partner drängt nicht" und "habe starke Alkoholprobleme, die Familie/der Partner drängt etwas".

Mit der Prädiktions-KFA konnten zwar "Prädiktionstypen" und "Prädiktionsantitypen" ermittelt werden, jedoch kann nicht, wie mittels der loglinearen Analyse möglich ist, sichtbar gemacht werden, durch welche Haupteffekte bzw. durch welche Interaktionseffekte unterschiedlicher Ordnung diese determiniert werden. Damit sind auch die Vergleichsmöglichkeiten der (Un-)Ähnlichkeiten der einzelnen "Prädiktionstypen" und "Prädiktionsantitypen" sehr begrenzt, was insbesondere für die lediglich tendenziell nachgewiesenen zutrifft. Als explorativen Hilfsmittel kann für diesen Zweck wiederum die Korrespondenzanalyse verwendet werden, somit ist eine weitergehende Beschreibung der Ergebnisse der KFA möglich.

8. Dritter Schritt: Grafische Darstellung der Ergebnisse der KFA

In diesem Abschnitt werden die im Abschnitt 7 nachgewiesenen, als auch die als nicht signifikant ermittelten, "Prädiktionstypen" und "Prädiktionsantitypen" graphisch abgebildet, so daß (Un-)Ähnlichkeiten zwischen diesen sichtbar werden. Eingabeinformation für die zweite Korrespondenzanalyse sind die absoluten Häufigkeiten, die aus der Kreuztabellierung mit der in der KFA verwendeteten zusammengesetzten Variable (neun Merkmale) und der Variable "Therapieteilnahme", hervorgegangen sind. Die Tabelle ist identisch mit Tabelle 4.

Tabelle 6: Eingabedaten der zweiten Korrespondenzanalyse

	nur Erst-gespräch	nur Vb.-gruppe	Therapie-teilnähme
Keine A.probl., Fam. drängt nicht	21	5	3
Keine A.probl., Fam. drängt etwas	9	4	0
Keine A.probl., Fam. drängt stark	7	10	0
Etwas A.probl., Fam. drängt nicht	18	10	5
Etwas A.probl., Fam. drängt etwas	12	9	1
Etwas A.probl, Fam. drängt stark	17	6	9
Starke A.probl., Fam. drängt nicht	57	41	47
Starke A.probL, Fam. drängt etwas	27	34	28
Starke A.probl., Fam. drängt stark	42	40	33

Tabelle 7: Chi-Quadrat-Statistik der zweiten Korrespondenzanalyse

	nur Erst-gespräch	nur Vb.-gruppe	Therapie-teilnähme
Keine A.probl., Fam. drängt nicht	6.63	2.28	2.48
Keine A.probl., Fam. drängt etwas	2.39	0.03	3.24
Keine A.probl., Fam. drängt stark	0.00	3.27	4.24
Etwas A.probl., Fam. drängt nicht	1.33	0.10	1.27
Etwas A.probl., Fam. drängt etwas	0.89	0.36	3.67
Etwas A.probl, Fam. drängt stark	0.01	0.27	0.21
Starke A.probl, Fam. drängt nicht	0.18	1.17	3.24
Starke A.probl, Fam. drängt etwas	2.71	0.60	1.51
Starke A.probl, Fam. drängt stark	0.71	0.06	0.65

Chi^2=43.5 mit df=16; total inertia=0.0861

Bei der Prädiktions-KFA konnte unter inferenzstatistischen Testbedingungen als einziger "Prädiktionstyp" die Merkmalskombination "keine Alkoholprobleme, die Familie/der Partner

drängt nicht" und "nur am Erstgespräch teilgenommen" nachgewiesen werden. Diese Zelle trägt auch bei der Korrespondenzanalyse am meisten zur Gesamtvariation des Modells bei, d.h. hier sind die größten Abweichungen von erwarteten und beobachteten Werten. Am wenigsten zur Gesamtvarianz trägt die Merkmalskombination "etwas Alkoholprobleme, die Familie/der Partner drängt stark" in allen Ausprägungen der zu beschreibenden Variablen bei - mit der Prädiktions-KFA konnte für diese drei Zellen nicht einmal tendenziell ein Zusammenhang zwischen den beiden Variablen nachgewiesen werden.

Wie auch schon in der ersten Korrespondenzanalyse ist das Gegensatzpaar auf der ersten Achse (erklärte Varianz: 72.0%) "Therapieteilnehmer" vs. "nur am Erstgespräch teilgenommen". Interpretieren wir die Variablenausprägungen anhand der (Un-)Ähnlichkeit ihrer Winkel zum Achsenkreuz, können wir den mit der Prädiktions-KFA ermittelten "Prädiktionstyp" auch in der Korrespondenzanalyse finden. Auch in der Korrespondenzanalyse kann das Merkmal "Therapieteilnehmer" am besten durch die Merkmalskombination "keine Alkoholprobleme, Familie drängt nicht" beschrieben werden; der Winkel (bezogen zum Achsenkreuz) zwischen diesen beiden Vektoren ist am kleinsten.

Abbildung 2: Legende

ERSTGS	nur Erstgespräch
VGRUPPE	nur Vorbereitungsgruppe
THERAPIE	Therapieteilnehmer
APNFDN	Keine A.probl., Fam. drängt nicht
APNFDE	Keine A.probl., Fam. drängt etwas
APNFDS	Keine A.probl., Fam. drängt stark
APEFDN	Etwas A.probl., Fam. drängt nicht
APEFDE	Etwas A.probl., Fam. drängt etwas
APEFDS	Etwas A.probl., Fam. drängt stark
APJFDN	Starke A.probl., Fam. drängt nicht
APJFDE	Starke A.probl., Fam. drängt etwas
APJFDS	Starke A.probl., Fam. drängt stark

Abbildung 2: Grafische Darstellung der zweiten Korrespondenzanalyse

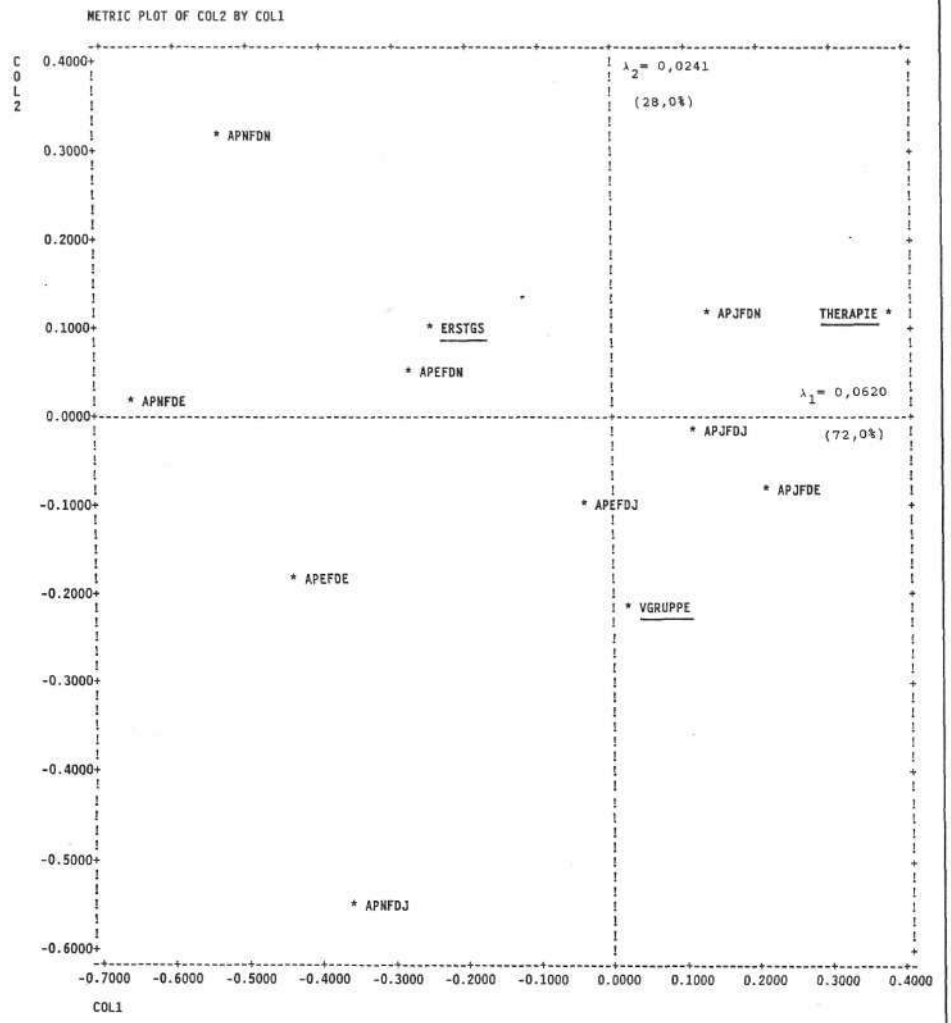


Tabelle 8: Numerische Darstellung der zweiten Korrespondenzanalyse

GENSTAT	MASS	SQCOR	INR	LOC1	QCOR1	INR1	LOC2	QCOR2	INR2
ERSTGS	0.416	1.000	0.341	-0.246	0.852	0.404	0.102	0.148	0.180
VGRUPPE	0.335	1.000	0.187	0.019	0.007	0.002	-0.219	0.993	0.663
THERAPIE	0.250	1.000	0.472	0.384	0.907	0.594	0.123	0.093	0.156
APNFDN	0.057	1.000	0.262	-0.541	0.746	0.271	0.316	0.254	0.237
APNFDE	0.026	1.000	0.130	-0.660	0.999	0.181	0.022	0.001	0.001
APNFDJ	0.034	1.000	0.173	-0.362	0.297	0.071	-0.557	0.703	0.434
APEFDN	0.065	1.000	0.062	-0.281	0.967	0.083	0.052	0.033	0.007
APEFDE	0.044	1.000	0.113	-0.437	0.853	0.134	-0.181	0.147	0.059
APEFDJ	0.083	1.000	0.011	-0.040	0.137	0.002	-0.100	0.863	0.035
APJFDN	0.287	1.000	0.105	0.134	0.565	0.083	0.117	0.435	0.164
APJFDE	0.176	1.000	0.111	0.215	0.853	0.131	-0.089	0.147	0.058
APJFDJ	0.228	1.000	0.033	0.109	0.960	0.043	-0.022	0.040	0.005

Neben dem einzigen (echten) "Prädiktionstyp" der Prädiktions-KFA können auch die anderen lediglich tendenziell (Lehmacher-Test ohne Alpha-Adjustierung) nachgewiesenen "Prädiktionstypen" mittels der Korrespondenzanalyse bestätigt werden - auch die Winkel zwischen diesen Vektoren sind relativ klein. Ferner können auch die "Prädiktionsantitypen", durch eine imaginäre Verbindung der jeweiligen Spalten- und Zeilenvariablenausprägungen, belegt werden - sie alle gehen relativ genau durch den Schwerpunkt oder anders ausgedrückt: Die Winkel zwischen diesen Vektoren sind sehr groß. Diese in beiden Dimensionen konträre Projektion auf den Achsen kann als negative Korrelation zwischen Zeilen- und Spaltenmerkmalen interpretiert werden.

Neben der Beschreibung der in der KFA gefundenen "Prädiktionstypen" und "Prädiktionsantitypen" können die Distanzen zwischen den Zeilen- sowie zwischen den Spaltenvariablen ebenfalls interpretiert werden; so z.B. die Ähnlichkeit der Merkmalskombinationen "keine Alkoholprobleme, die Familie/der Partner drängt etwas" und "etwas Alkoholprobleme, die Familie/ der Partner drängt nicht".

9. Fazit

Ziel dieses Beitrags war es, komplementäre Verwendungsmöglichkeiten der Korrespondenzanalyse und der KFA sowie Ähnlichkeiten dieser Verfahren aufzuzeigen. Hierbei konnte gezeigt werden, daß mit der (Prädiktions-)KFA statistisch signifikante "Typen" und "Antitypen" ermittelt werden können, wobei jedoch die Anzahl der zu verwendenden Variablen eng limitiert ist. Aus dieser Beschränkung folgt, daß die Auswahl der möglichen Variablen nicht mit dem Verfahren selbst erfolgen kann, sondern daß diese bereits bekannt sein müssen. Wenn dies nicht der Fall ist, können die relevanten Variablen vor der inferenzstatistischen Analyse mit einem explorativen Analyseverfahren ermittelt werden.

Bei der konfirmatorischen Datenanalyse entscheidet über "Signifikanz" und "Nicht-Signifikanz" oft nur die zweite Stelle hinter dem Komma, so daß die Testentscheidung häufig willkürlich erscheint. Um diese Differenzen nicht übermäßig zu bewerten, können mittels der Korrespondenzanalyse alle Abweichungen in den Daten grafisch (und numerisch) dargestellt und interpretiert werden. Ein derartiges Vorgehen bezüglich der komplementären Verwendung der Korrespondenzanalyse und log-linearer Modelle schlagen z.B. ***Goodman*** (1986) und ***van der Heijden et al.*** (1989) vor.

Die Korrespondenzanalyse eignet sich zum Aufspüren und Beschreiben nicht-trivialer Strukturen in mehrdimensionalen Daten, nicht aber zur Prüfung der Zusammenhänge auf Signifikanz. Sollen Aussagen nicht nur auf der beschreibenden Ebene, sondern auch auf der erklärenden gemacht werden, so muß ein konfirmatorisches Verfahren (z.B. die KFA) verwendet werden. Da sowohl die explorative Beschreibung von Zusammenhängen als auch das Testen auf Signifikanz wichtige Ziele der empirischen sozialwissenschaftlichen Forschung sind, bleibt die Frage: Warum nicht zwei Verfahren komplementär verwenden?

Mit diesem Beitrag wurde ein derartiger Vorschlag zur komplementären Verwendung von konfirmatorischen und explorativen Verfahren gemacht, hier dargestellt am Beispiel der Korrespondenzanalyse und der (Prädiktions-)KFA. Konfirmatorische und explorative Analyseverfahren sind u.E. auch keine Gegensätze, wie implizit und teilweise auch explizit immer wieder behauptet wird, sondern Ergänzungen. Nicht zu Unrecht fordert ***Tukey*** (1980) im Titel eines Beitrages: "We Need Both Exploratory and Confirmatory".

Literatur

J.-P. *Benzécri et collaborateurs* (1973):
L'Analyse des Donn6es II. L'Analyse des Correspondances. Paris: Dunod.

Jörg *Blasius* (1987):
Einstellungen zur Hamburger Innenstadt. Eine Auswertung mit Hilfe der Korrespondenzanalyse.
ZA-Information, 21, S. 29-51.

Jörg *Blasius* and Harald *Rohlinger* (1989):
KORRES - A Program for Multivariate Analysis of Categorical Data from Contingency Tables.
In: SAS Institute Inc. (eds.): SEUGI'89. Proceedings of the SAS Users Group International Conference.
Gary, N.C.: SAS Institute Inc., S. 98-117.

Jörg *Blasius* und Joachim *Winkler* (1989):
Gibt es die feinen Unterschiede? Eine empirische Überprüfung der Bourdieuschen Theorie.
Kölner Zeitschrift für Soziologie und Sozialpsychologie, 41, S. 72-94.

Pierre *Bourdieu* (1982):
Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft. Frankfurt/M.: Suhrkamp.

Pierre *Bourdieu* (1988):
Homo Academicus. Frankfurt/M.: Suhrkamp.

Jens *Dangschat* and Jörg *Blasius* (1987):
Social and Spatial Disparities in Warsaw in 1978:
An Application of Correspondence Analysis to a 'Socialist' City. Urban Studies, 24, S. 173-191.

Alexander *von Eye* (1990):
Introduction to Configural Frequency Analysis. New York: Cambridge University Press.

Leo A. *Goodman* (1986):
Some Useful Extensions of the Usual Correspondence Analysis Approach and the Usual Log-Linear Models Approach in the Analysis of Contingency Tables.
International Statistical Review, 54, S. 243-309.

Michael *Greenacre* (1984):
Theory and Applications of Correspondence Analysis. London: Academic.

Michael *Greenacre* (1989):
The Carroll-Green-Schaffer Scaling in Correspondence Analysis: A Theoretical and Empirical Appraisal.
Journal of Marketing Research, 26, S. 358-365.

Michael *Greenacre* (1990):
SimCA 2.0. Pretoria.

Michael *Greenacre* and *TrevorHastie* (1987):
The Geometrie Interpretation of Correspondence Analysis.
Journal of the American Statistical Association, 82, S. 437-447.

Joachim ***Krauth*** (1983):
Typenanalyse. In: Jürgen ***Bredenkamp*** und Hubert ***Feger*** (Hrsg.): Forschungsmethoden der Psychologie, Bd. 4, Enzyklopädie der Psychologie. Göttingen: Hogrefe, S. 440-496.

Joachim ***Krauth*** und Gustav **A.** ***Lienert*** (1973):
KFA. Die Konfigurationsfrequenzanalyse und ihre Anwendung in Psychologie und Medizin. Freiburg: Alber.

Erwin ***Lautsch*** und Stefan ***von Weber*** (1990):
Die Konfigurationsfrequenzanalyse (KFA) - Methoden und Anwendungen. Berlin: Volk und Wissen.

Ludovic ***Lebart,*** Alain ***Morineau*** and Kenneth M. ***Warwick*** (1984):
Multivariate Statistical Analysis. Correspondence Analysis and Related Techniques for Large Matrices. New York: Wiley.

Gustav A. ***Lienert*** (1969):
Die "Konfigurationsfrequenzanalyse" als Klassifikationsmittel in der klinischen Psychologie. In: Martin ***Irle*** (Hrsg.): Bericht 26. Kongreß der Deutschen Gesellschaft für Psychologie in Tübingen 1968. Göttingen: Hogrefe, S. 244-253.

Gustav A. ***Lienert*** (1988):
Angewandte Konfigurationsfrequenzanalyse. Frankfurt/M.: Athenäum.

Viktor ***Thiessen*** und Harald ***Rohlinger*** (1988):
Die Verteilung von Aufgaben und Pflichten im ehelichen Haushalt.
Kölner Zeitschrift für Soziologie und Sozialpsychologie, 40, S. 640-658.

John ***W.Tukey*** (1980):
We Need Both Exploratory and Confirmatory.
The American Statistician, 34, S. 23-25.

Peter G.M. ***van der*** *Heijden* and Jan ***de Leeuw*** (1985):
Correspondence Analysis Used Complementary to Loglinear Analysis.
Psychometrica, 50, S. 429-447.

Peter G.M. ***van der*** *Heijden,* Antoine ***de Falguerolles*** and Jan ***de Leeuw*** (1989):
A Combined Approach to Contingency Table Analysis Using Correspondence Analysis and Log-Linear Analysis.
Applied Statistics, 38, S. 249-292.

Nun kann auch die soziologische Methodik zusammenwachsen

von Michael Häder

Nachdem W. ***Brandt*** das Zusammenwachsen Deutschlands und R. ***Lepsius*** das Zusammenwachsen der Soziologie beschworen haben, ist es an der Zeit, auch etwas für die Vereinigung der Methodik und die Verständigung der Methodiker zu tun.

Die kritische Aufarbeitung von nun abgeschlossenen 26 Jahren institutioneller Soziologie in der DDR ist sowohl Voraussetzung für das zukünftige Selbstverständnis der ehemaligen DDR-Soziologen als auch für die anstehende Vereinigung mit den Methodikern der "alten" Bundesrepublik. Die Aufgabe der eigenen Vergangenheitsbewältigung muß vor allem von den Soziologen aus der ehemaligen DDR selbst geleistet werden. Es wird sich dabei um einen komplizierten und langfristigen Vorgang handeln, in dessen Rahmen vor allem die Fragen nach der (Mit-)Verantwortung der Soziologen für die Krise in der DDR und für die Apologetik des Systems zu beantworten sind. Eine solche Aufarbeitung kann an dieser Stelle und in diesem Rahmen nicht geleistet werden. Einige Gedanken zur Methodik wurden indes auf der Jahrestagung der ASI im Oktober 1990 in Berlin zur gegenseitigen Selbstverständigung vorgestellt und sollen an dieser Stelle kurz referiert werden.

Nachdem 1964 das Politbüro beim Zentralkommitee der SED die Zulassung von marxistisch-leninistischer Soziologie in der DDR beschlossen hatte, entwickelten sich rasch auch empirische Studien. Kurt *Braunreuther* in Berlin und Robert *Schulz* in Leipzig als die Wegbereiter der Soziologie in der DDR konnten jeder für sich bereits eine starke empirische Orientierung in Anspruch nehmen. Seit Beginn der empirisch-sozialwissenschaftlichen Arbeiten in der DDR bestand für diese ein Konfliktpotential gegenüber den Interessen der Parteiführung. Erste Befragungen wurden von der Staatssicherheit noch vor deren Auswertung beschlagnahmt. Dies zeigt, daß sich die empirisch-soziologische Forschung in der DDR von Anfang an einer strengen Aufsicht sicher gewesen sein mußte. Mit der Institutionalisierung der DDR-Soziologie hat sich auch die Kontrolle über deren empirische Arbeit weiter etabliert. Eine nicht nur für Außenstehende kompliziert zu durchschauende und dazu strenge Genehmigungsprozedur für repräsentative Bevölkerungsbefragungen, die sich über Jahre hinausziehen konnte, machte kurzfristige Erhebungen unmöglich und erschwerte damit solche Studien erheblich. Auch die Ergebnisse anderer Untersuchungen (vor allem in Betrieben) unterlagen strengen Auflagen zur Genehmigung und danach zur Geheimhaltung.

In seinem Überblick über die universitäre Soziologie in der DDR hat *Lepsius* (1990) auch für den akademischen Sektor zutreffende Aussagen gemacht. Die Gültigkeit einer Reihe von

Feststellungen für die in der ehemaligen DDR betriebene soziologische Methodik soll hier noch einmal hervorgehoben und unterstrichen werden. Dies trifft z.B. für das in der "alten" Bundesrepublik "reichere" und ausdifferenziertere, arbeitsteiligerer betriebene Niveau der methodischen Arbeit zu. Von den 700 bis 800 Soziologen der ehemaligen DDR dürften sich derzeit nicht einmal fünf Prozent als Methodiker verstehen. Der bis Mitte 1989 in der DDR bestandene Problemrat Methodologie/Methodik umfaßte zuletzt nicht mehr als 40 aktive Mitglieder. Die Schwierigkeiten, eigene Forschungsergebnisse zu publizieren, trafen für die empirisch arbeitenden Sozialforscher in besonderem Maße zu. Dies zum einen deshalb, weil es mit soziologischen Fachzeitschriften ohnehin schlecht bestellt war, und zum anderen, weü gerade methodische Forschungsergebnisse vielfach mit geheimzuhaltenen Zahlenmaterial verbunden waren.

So trifft auch noch eine weitere Aussage *Lepsius'* über die universitäre Soziologie gerade auch auf die Methodik der gesamten DDR-Soziologie voll zu: Die empirisch-soziologische Forschung in der DDR war weitaus umfänglicher als dies aus den Publikationen zu ersehen ist. Schließlich hat auch eine andere seiner Feststellungen Konsequenzen für die Methodik; für die DDR-Soziologie waren makrosoziologische Fragestellungen bzw. die institutionelle Struktur der DDR als Forschungsthemen weitgehend tabuisiert. Die Folge daraus ist ein Defizit auf dem entsprechenden Teilgebiet der soziologischen Methodik. Schließlich trifft es ebenfalls zu, daß die Methodiker in der DDR eine gründlichere Rezeption westlicher Literatur betreiben konnten und betrieben haben, soweit die erforderliche Literatur erreichbar war - als andere Zweigsoziologen. Dies gilt selbst für die studentische Methodik-Ausbildung.

Die Arbeit der Methodiker in der DDR läßt sich gut verdeutlichen anhand der Tätigkeit des Problemrates Methodologie/Methodik beim Wissenschaftlichen Rat für die soziologische Forschung in der DDR. Dieser Problemrat war die Vereinigung der DDR-Methodiker unterstellt, d.h. rechenschaftspflichtig dem Wissenschaftlichen Rat für die Soziologische Forschung in der DDR, und damit "das Gegenstück" zur Methodensektion der Deutschen Gesellschaft für Soziologie der Bundesrepublik. Der Problemrat war jedoch offen für alle Interessenten und umfaßte in den knapp 20 Jahren seines Bestehens durchschnittlich etwa 40 bis 45 Kollegen aus ca. 20 Einrichtungen der ehemaligen DDR. Jährlich fanden immerhin drei eintägige und eine zweitägige Veranstaltung statt. In einer doch sehr kollegialen und offenen Weise haben die Methodiker versucht, mit der für die damalige DDR typischen Forschungs Situation, die gekennzeichnet war durch Geheimhaltung, Publikationsprobleme, Devisenmangel, Nöten bei der Literaturbeschaffung, Austausch von illegal beschaffter Soft-Ware, Defizit an soziologischer Theorie usw., fertig zu werden. Obwohl gerade die Methodiker von Restriktionen in ihrer Forschungstätigkeit betroffen waren, kam von ihnen, wie auch von den meisten anderen DDR-Soziologen, kein ernsthafter Protest gegen die beschränkten Arbeitsbedingungen und schon gar nicht gegen die gesellschaftlichen Verhältnisse zustande. So war die Arbeit des Pro-

blemrates vor allem gekennzeichnet durch das Bemühen, die unliebsamen Arbeitsbedingungen gemeinsam zu lindern bzw. zu kompensieren und nicht so sehr durch das Bestreben, vereint eine (grundsätzliche) Veränderung dieser Bedingungen zu erreichen.

Weiteres allgemeines Kennzeichen der DDR-Soziolgie war die Personifizierung von Forschungslinien. Es fand kaum ein wissenschaftlicher Disput zu methodologischen und methodischen Grundfragen statt, wie er für eine einheitliche, als marxistisch-leninistisch bezeichnete Soziologie auch gar nicht vorgesehen war. So standen vielmehr einzelne Vertreter als Oberhäupter den einzelnen soziologischen Teildisziplinen "vor". Dies galt auch für die Methodologie/Methodik der soziologischen Forschung.

Um so deutlicher sich jedoch für die Soziologen auf der einen Seite die politischen Grenzen für ihre Arbeit spürbar machten, desto mehr wurde nach systeminternen Kompensationsmöglichkeiten gesucht. Diese fanden, mitunter als relativ große, ideologiefreie Nischen gerade die Methodiker in ihrer Arbeit. So hat sich besonders in den letzten Jahren in der DDR-Soziologie in harmonischer Koexistenz eine sehr differenzierte sowohl qualitativ als auch quantitativ orientierte Erhebungs-Methodik entwickelt. Daneben bot die mathematisch-statistische Datenanalyse ein weiteres Feld, auf dem breiter gearbeitet werden konnte. Beispielsweise befaßte sich der Problemrat bereits 1978 erstmals auf einer Tagung mit der Anwendung von Pfadanalysen. Immer wieder waren originell ausgearbeitete Erhebungsstrategien, methodologische Fragen, Ansätze für erweiterte Datenanalysen und, soweit möglich, auch inhaltliche Ergebnisse der empirischen Untersuchungen Gegenstand von internen Fachdiskussionen der Methodiker. So hat sich in der Soziologie der damaligen DDR eine Methodik entwickelt, die sich in ihren Forschungen konzentrierte auf:

- mathematisch-statistische Datenanalyse,
- originelle erhebungsmethodische Forschungsdesigns, die sich an Individuen richteten,
- auf die damit zusammenhängenden methodologischen Probleme, die sogar über eine "Standardmethodologie" *(Lepsius)* hinausgegangen sein dürften und auf
- Indikatorenentwicklungen zu partiellen Sachverhalten, eingeschlossen die Bearbeitung der entsprechenden Güteprobleme.

Diese Vielfalt ist um so beachtlicher, wenn man das doch geringe Potential an Methodikern und deren komplizierten Arbeitsbedingungen in Rechnung stellt. Einen Einblick in die von der DDR-Soziologie betriebene methodische Arbeit soll folgende Themenübersicht geben. Eine solche Aufzählung erhebt keinesfalls den Anspruch auf Vollständigkeit. Hervorzuheben sind:

- Arbeiten an der Standardisierung von Fragebatterien, insbesondere für den Demographie-Teil, in Bevölkerungsbefragungen der damaligen DDR (Institut für Soziologie und Sozialpolitik, Berlin).

- Der Einsatz von **komplexen Methodiken** in einer von der Akademie der Pädagogischen Wissenschaften in Berlin 1983 begonnenen Schüler-Panel-Untersuchung zu Berufslaufbahnen: Über einen längeren Zeitraum wurden Schulklassen in verschiedenen Orten mittels standardisierter Befragung, Gruppendiskussion, Beobachtung, Tiefeninterviews, Dokumentenanalyse und Elternbefragungen begleitend empirisch beobachtet. Schüler-Intervallstudien mit soziologischem Anspruch wurden z.B. auch vom Zentralinstitut für Jugendforschung in Leipzig und von der Pädagogischen Hochschule in Erfurt angelegt.
- Die **Intervallstudien,** die z.B. bei Sportlern (Thema: politisch-ideologische Haltung von Nachwuchs-Fußballern!), Hochschulabsolventen und im Dienstleistungsbereich unter den Aspekten verschiedener Forschungseinrichtungen betrieben wurden (Deutsche Hochschule für Körperkultur und Sport, Leipzig, Zentralinstitut für Hochschulbildung, Berlin, und Institut für Soziologie und Sozialpolitik, Berlin.
- Die **teilnehmenden Beobachtungen** in automatisierten Produktionsbereichen zur Prozeßforschung; längerfristig arbeiteten Soziologen in den Untersuchungsbereichen mit und analysierten hier vor allem spezifische Gruppenprozesse (Technische Universität Dresden).
- Die ebenfalls mit arbeits- bzw. betriebssoziologischem Anspruch angelegte repräsentative **Vergleichsstudie** DDR-Sowjetunion zur Einführung neuer Technologien in der Industrie (Akademie für Gesellschaftswissenschaften beim ZK der SED, Berlin).
- Die Adaption von Mikrosimulationsmodellen im Rahmen der Ausarbeitung von Prognosemethoden für Bedürfnisse (Institut für Soziologie und Sozialpolitik, Berlin).
- Die methodologischen **Studien** zur Rolle des individuellen sozialen Subjekts in empirischen Erhebungen, in deren Mittelpunkt das Vermögen des einzelnen als Auskunftsperson über Soziales aufzutreten stand (Karl-Marx-Universität Leipzig, Institut für Soziologie und Sozialpolitik, Berlin).
- **Komplexe Fallstudien** an einem wissenschaftlichen Forschungsinstitut, die von der Beobachtung des Forschungsalltags ausgingen und bis hin zur Betrachtung der Freizeitgestaltung und der familialen Situation von ehemaligen DDR-Spitzenwissenschaftlern reichten (Karl-Marx-Universität, Leipzig).
- Die computergestützte Variante einer **Entscheidungs- und Orientierungshilfe** (EOH) für die Auswahl mathematisch-statistischer Auswertungsmethoden in der Soziologie (Institut für Soziologie und Sozialpolitik, Berlin).
- Die territoriale soziologische Studie zu **Kulturbedürfnissen** in der Stadt Jena (Friedrich-Schiller-Universität, Jena).
- Die Repräsentative Befragung von Ärzten zur **medizinischen Betreuung** im Raum Magdeburg (Institut für Planung und Organisation des Gesundheitswesens, Magdeburg).
- **Militärsoziologische empirische Untersuchungen** unter Manöverbedingungen, für die eigens eine "Befragungsmaschine" entwickelt wurde (Politische Hauptverwaltung der NVA, Strausberg).

Bemerkenswert ist das breite Spektrum, in dem empirisch-soziologische Studien betrieben wurden. Nahezu alle gesellschaftlichen Bereiche der ehemaligen DDR waren Gegenstand konkreter Sozialforschungen. So haben sich nicht nur die (wenigen) soziologischen Lehrstühle, sondern auch zahlreiche andere Universitäten und akademische Institute mit empirischen Studien befaßt. Zu nennen sind beispielweise die Hochschule für Verkehrswesen, Dresden, die Hochschule für Architektur und Bauwesen, Weimar, die Wilhelm-Pieck-Universität, Rostock, die Charité, Berlin, die Technische Hochschule Leuna-Merseburg u.a. Bekannt wurden auch größere Untersuchungen des Zentralinstituts für Jugendforschung, Leipzig, des Zentralinstituts für Ernährung, Potsdam-Rehbrücke, des Zentralinstituts für Arbeitswissenschaften, Dresden, der Bauakademie, Berlin, des Rundfunks der DDR u.a. Das Institut für Soziologie und Sozialpolitik der ehemaligen Akademie der Wissenschaften der DDR und die Soziologie als Hauptfach ausbildenden Universitäten sowie die ehemalige Akademie für Gesellschaftswissenschaften beim Zentralkommitee der SED haben die wohl umfangreichste empirische Arbeit betrieben.

Der Freiraum, der den Methodikern zur Verfügung stand, ermöglichte immerhin kritische Bestandsaufnahmen und schaffte damit eine Voraussetzung für die wissenschaftliche Arbeit. Gegenüber anderen soziologischen Disziplinen, deren Vorgehen stark monolithischen Charakter trug, war die soziologische Methodik dazu in der Lage, gegenseitige wissenschaftliche Kritik an den Ergebnissen ihrer Arbeit zu äußern. In einer methodenkritischen Aufarbeitung von über 40 zu dieser Zeit mitunter vertraulicher Forschungsberichte verschiedener Institutionen (vgl. *Hader,* 1987) ergab sich zusammengefaßt:

- Es fehlen Erkenntnisse zum Zusammenhang zwischen 1. den vom Subjekt gewünschten Informationen, 2. der Leistungsfähigkeit der Methodik und 3. den Erkenntnisbesonderheiten der Objekte soziologischer Forschung sowie zu jedem dieser drei Elemente.
- Die Leistungsfähigkeit einzelner Methoden ist zu wenig bekannt, weshalb es zu Fehlinterpretationen (Meinungen über Tatsachen werden z.B. in Form von Faktenaussagen zu den Tatsachen selbst uminterpretiert.) der gewonnenen empirischen Informationen kommt.
- Die vom Subjekt gewünschten Erkenntnisse sind zu wenig theoretisch vorstrukturiert. Es finden sich verschwommen formulierte und zumindest vordergründig wenig soziologisch anmutende Themen, die weitergeführt in unkonkret ausgedrückte Problemstellungen den Erkenntnisgewinn der gesamten Untersuchung vermindern.
- Wesentliche Eigenschaften der soziologischen Untersuchungsobjekte, vor allem deren Erscheinungsweise, wurden zu wenig zur Grundlage für die Ausarbeitung der Methodik gemacht. So mangelte es vielen empirisch-soziologischen Erhebungen noch an Zielgerichtetheit und Systematik.

Die gesellschaftlichen Veränderungen vom Herbst 1989 brachten völlig neue Bedingungen für die Arbeit der Soziologie-Methodiker mit sich: Die Genehmigungspflicht für repräsentative Untersuchungen entfiel, die Geheimhaltungsvorschriften wurden außer Kraft gesetzt, die thematische Ausrichtung der Forschungen durch den Wissenschaftlichen Rat und damit letztendlich durch die SED-Führung hörte auf, das Kontaktverbot zu Kollegen aus westlichen Ländern wurde aufgehoben, die Genehmigungspflicht für Publikationen wurde gestrichen, Service-Einrichtungen für Umfragen entstanden u.a. Damit bestand seit dem November 1989 die prinzipielle Möglichkeit, die gesellschaftlichen Veränderungen begleitend empirisch-soziologisch zu beobachten. Diese Chance wurde von den nun ehemals DDR-Soziologen mehr oder weniger genutzt. Bis dahin unvorstellbar gewesene Befragungen zu Themen wie die deutsche Vereinigung, zur Sympathie und Antipathie gegenüber Politikern und Parteien, zur politischen und wirtschaftlichen Situation, zum Zusammenleben mit Ausländern u.a. fanden statt. Von den Soziologen der Leipziger Universität wurden z.B. auch die Teilnehmer der wöchentlichen Leipziger Montagsdemonstration regelmäßig befragt. So sind in dieser einmaligen Zeit der gesellschaftlichen Veränderungen dank der Initiative der vor allem methodisch arbeitenden Kollegen einige wertvolle (wenn auch sicherlich immer noch viel zu wenige) empirische Zeitzeugnisse entstanden, die die Situation dieses Abschnittes der deutschen Geschichte dokumentieren.

Während *Lepsius* eine Neukonstituierung der Soziologie in der ehemaligen DDR für nötig hält, kann deren Methodik doch an einige bisherige Arbeitslinien anknüpfen. Ihre Fähigkeit dazu wird sie bei der Lösung einer Reihe dringender Probleme unter Beweis stellen müssen. So geht es derzeit um:

- Die Ausarbeitung von Standarddemographien, die den sozialstrukturellen Bedingungen im gesamten Bundesgebiet gerecht werden. Insbesondere die Fachschul- und Facharbeiterqualifikationen der DDR können nicht ohne weiteres mit entsprechenden Abschlüssen aus der Bundesrepublik in Beziehung gesetzt werden. Dies gilt auf der anderen Seite für den Beamtenstatus ebenfalls, dem eine Entsprechung in der DDR fehlte. Analoge Demographie-Indikatoren sind jedoch die Voraussetzung für methodisch gesicherte Ost-West-Vergleiche.
- Die in der Bundesrepublik als Zeitreihen vorliegenden Daten wie z.B. Ingelhart-Index, Rechts-Links-Skala, Oben-Unten-Skala, "Sonntagsfrage" usw. sollten auch auf dem Gebiet der ostdeutschen Länder möglichst schnell angewandt werden. Dazu sind methodische Tests, insbesondere zur sprachlichen Validierung, dringende Voraussetzung.
- Die Datenarchivierung, in der bisherigen Bundesrepublik vom Zentralarchiv koordiniert, muß auf das Gebiet der ehemaligen DDR ausgedehnt werden. Hier vorliegende, bisher als streng geheim behandelte, und - wenn überhaupt - dezentral gelagerte empirische Datensätze sind aus methodischer Sicht systematisch aufzuarbeiten und der wissenschaftlichen Öffentlichkeit zugänglich zu machen. - Das Niveau der kontinuierlichen Datenerhebung, welches in der Bundesrepublik z.B. vom ALLBUS dem Sozio-Ökonomischen Panel, dem

Wohlfahrtssurvey, dem ISSP u.a. gesichert wird, gilt es auch in den neuen Bundesländern zu erreichen. Dabei dürfte eine einfache Ausdehnung der Stichprobe auf das ehemalige DDR-Gebiet nicht ausreichen. Es müssen zusätzlich geeignete Formen gefunden werden, um die derzeit auf dem Gebiet der ehemaligen DDR mit einer hohen Dynamik ablaufenden sozialen Prozesse kontinuierlich zu beobachten, z.B. sollten Wiederholungsuntersuchungen in relativ kurzen Abständen (ca. 6 Monate) für die Länder Ostdeutschlands theoretisch konzipiert und methodisch umgesetzt werden.

Die mit der Integration der ehemaligen DDR sich vollziehenden sozialen Prozesse könnten zum Modell für eine europäische Integration werden. Derzeit im Osten Deutschlands auszumachende soziale Anpassungsprozesse sollten unter diesem Aspekt für drei bis vier Jahre kontinuierlich und besonders intensiv empirisch beobachtet werden. Es entstehen mit der Konstituierung der Länder und mit dem gleichzeitigen Ende der DDR neue Voraussetzungen für das Selbstverständnis der hier lebenden Menschen. So könnten sich aus der gegenwärtigen Identitätssuche in der ehemaligen DDR interessante Aussagen ableiten, die für die anstehende europäische Integration von besonderem Interesse sein könnten.

Literatur

Häder, Michael.,
Zur Rezeption einer speziellen Methodologie für empirisch-soziologische Erhebungen, Diss. B (Habilitationsschrift), Humboldt-Universität zu Berlin, 1987.

Lepsius, Rainer M„
Zur Lage der Soziologie an den Universitäten der DDR, Kölner Zeitschrift für Soziologie und Sozialpsychologie, 42 (1990) 2, S. 313 - 323.

Michael Häder
Institut für Soziologie und Sozialpolitik
Otto-Nuschke-Str. 10/11
O -1086 Berlin

Eine Wiederbesichtigung des Forschungsbetriebs "Soziologie" in der UdSSR[1]

von Wolfgang Teckenberg

Mit diesem Aufsatz wird an meinen Artikel in der ZA-Information Nr. 23 von Ende 1988 angeknüpft. Daher handelt es sich hier vor allem um die Darstellung "neuerer" Entwicklungen der sowjetischen Soziologie, und es werden drei Thesen über offensichtliche Stagnationstendenzen entwickelt. Diese beziehen sich vor allem auf die Strukturen des Wissenschaftsbetriebs, nicht auf die Inhalte, die im ganzen insofern positiver bewertet werden können, als heute mannigfaltigere Themenbereiche angesprochen werden, die vor Jahren noch im Bereich der Tabuzone lagen. Auch geht der Artikel nicht auf die Lage der Lehre des Fachs ein, die allerdings in nennenswertem Maße bisher nur an der Moskauer Universität betrieben wird. Ich beziehe mich also im wesentlichen auf die Forschung, wie sie vor allem im Rahmen der Akademie der Wissenschaften und an einigen aufzuführenden Forschungsinstituten stattfindet.

"Neu" scheint gemeinhin die Entwicklung ab März 1985, dem Amtsantritt *Gorbatschows*, zu sein. Doch wie gerade Soziologen wissen, verändern sich Strukturen und die einmal auf eine Parteilinie eingeschworene Institution "Soziologie" selten sogleich mit dem Regierungswechsel.

Die im übrigen sicherlich positiv zu bewertende Absicherung der Soziologie durch das offizielle Placet des ZK-Beschlusses vom 12. Juni 1988 "Über die Stärkung der Rolle der marxistisch-lenistischen Soziologie für die Lösung kritischer Probleme der sowjetischen Gesellschaft" habe ich bereits ausführlicher in oben erwähnter Ausgabe dargestellt. Dort findet sich auch eine Einschätzung der Situation in der Lehre des Fachs und ein kurzer Aufriß der nun möglich gewordenen Forschungen sowie eine Qualitätsbeurteilung derselben.

---

1 Überarbeitete Fassung eines Papers für die Ad-hoc-Gruppe "Osteuropa" auf dem 25. Deutschen Soziologentag in Frankfurt am 19.10.1990.
Siehe zum Thema auch meine Aufsätze:
(1984) Organisation und Funktion sowjetischer Soziologie. Erforschung oder Leitung der Gesellschaft? In: Beiträge zur Konfliktforschung (14) Nr. 4, S.35-58.
(1988) "Perestrojka" in der sowjetischen Soziologie? Was man hört, was man sieht. In: ZA-Information Nr. 23 (November), S. 92-104.
Für wertvolle Hinweise danke ich Frau Natalija *Sorkaja* aus Moskau, die ich anläßlich ihres Forschungsaufenthalts am Zentralarchiv für empirische Sozialforschung in Köln sprechen konnte.

Hier will ich im wesentlichen nur drei Thesen zur Inertia des Forschungsbetriebs "Soziologie" unter sowjetischen Organisationsbedingungen vorstellen. Dies setzt allerdings voraus, daß ich zunächst einige Fakten referiere, die dem einen "neu", dem anderen bekannt vorkommen dürften. Dabei klammere ich bewußt gewisse Sonderentwicklungen in den baltischen Republiken aus, die sich - bezogen auf Studien in diesen Gebieten - schon vor 1985 abzeichneten.[2]

Als "etabliert" könnte man folgende Institutionen bezeichnen, wenn auch die Direktoren dieser Institute häufig wechselten und seit 1986 auch in ranghohe Stellen der Sowjetischen Soziologischen Vereinigung (siehe unten, These 2), die 1958 gegründet wurde, früher als kritisch bis dissidierend eingestufte Soziologen einrückten:

- Das Institut für Soziologie (unter diesem Namen seit 1988) der Akademie der Wissenschaften in Moskau und dessen Direktor, Vladimir A. ***Jadov,*** der im übrigen, wie alle weiteren Institutsleiter, Vizepräsident der Sowjetischen Soziologischen Vereinigung ist. Das Institut gilt als **die** zentrale Einrichtung für Grundlagenforschung im Bereich der Soziologie, und ihm obliegt auch die Koordination kleinerer Forschungsgruppen im Lande. Die Ergebnisse der Arbeit werden zunächst nicht für den allgemeinen Büchermarkt, sondern in sogenannten "Rotaprint-Ausgaben" mit kleiner Auflage (ca. 500 Exemplare) verbreitet. Dies reicht nicht einmal für alle Mitglieder der Sowjetischen Soziologen-Vereinigung. Neuerdings wurde ein Plan dieser Publikationen für 1990 veröffentlicht[3], der 35 Titel aufführt.

- Die soziologische Sektion beim **Institut für Ethnographie** der Akademie der Wissenschaften in Moskau und deren Direktor, Ovsej I. ***Schkaratan.***

- Die mit Sozialstrukturanalysen der UdSSR im weitesten Sinne befaßte Sektion am **Institut für Internationale Arbeiterbewegung** (Moskau) mit bekannten Forschern, wie E. V. ***Klopov,*** L. A. ***Gordon***, A. K. ***Nazitnova*** und anderen.

- Genauere Angaben zu dem sich in Leningrad befindenden **Institut für sozio-ökonomische Probleme** der Akademie der Wissenschaften fehlen mir z. Z.; generell ist jedoch eine Konzentration auf Moskau überdeutlich, die auch dazu führte, daß früher in Leningrader

2 So konnte der bekannte estnische Soziologe Mikk Ch. *Titma* sogar stellvertretender Parteisekretär der Republik werden. Wie man hört, hat er aber seit dem Sommer 1990 der Politik wieder den Rücken gekehrt und steht im Begriff, ein Forschungsstipendium nach Stanford anzunehmen.

3 Siehe Soziologitscheskie Issledovanija (künftig S. I.) 1990, Nr. 8, S. 157.

Nischen überwinternde Soziologen wieder nach Moskau gingen, so auch der genannte **V. A. *Jadov.***

Auch sowjetischen Fachleuten in Sachen "Wissenschaftsorganisation" war aufgefallen, daß es oft leichter ist, ein neues Institut zu gründen als ein altes umzustrukturieren. Probleme gibt es natürlich dann bei dem Bemühen, dafür zahlungskräftige Klientele zu gewinnen. Dieses Problem schien bei der Neugründung folgender Institute gelöst:

- Der volle Name drückt es aus: "**Allunionszentrum zur Erforschung der öffentlichen Meinung zu sozial-ökonomischen Fragen bei den Gewerkschaften und dem Staatskomitee für Arbeit**", das ich hinfort aber als "Institut für Meinungsforschung" oder nach der Direktorin, Präsidentin der Sowjetischen Soziologischen Vereinigung und *Gorbatschow* Beraterin Tatjana *Saslawskaja* als "Saslawskaja Institut" bezeichnen werde.[4] Wie gesagt, die Finanzprobleme "schienen" gelöst. Als Einheitsgewerkschaft ist die sowjetische "Arbeitervertretung" zwar finanzstark, jedoch vermutlich mehr an der Wahrung des sozialistischen Gesellschaftsvertrags interessiert als die ausgesprochen reformorientierte Streiterin für die Perestrojka, Tatjana *Saslawskaja*. Jedenfalls fließt zunächst aus der Gewerkschaftsquelle kein Geld mehr. Versuche, sich als Auftragsforschungsinstitut für das neu gewählte Parlament, den Obersten Sowjet, zu verdingen, scheiterten an dessen leeren Kassen. Der stellvertretende Direktor, auf dessen ansehnliche Reputation im Bereich der empirischen Sozialforschung man ebenfalls bauen konnte, Boris *Gruschin*, hat das Institut bereits verlassen. Frau *Saslawskaja*, die bei ihrer Besichtigung der Soziologie in der Bundesrepublik letztes Jahr vor allem bei Frau *Noelle-Neumann* verweilte und weniger an soziologischer Grundlagenforschung interessiert war, ist sehr im Rahmen der direkten Politikberatung engagiert.[5] Daher obliegt heute die eigentliche Umfrageaktivität einer durchaus befähigten und mit den Problemen landesweit repräsentativer Samples vertrauten Gruppe um den früher verfemten Jurij *Levada*.[6] Auf Möglichkeiten, die gegenwärtigen Finanzierungsprobleme des Instituts zu umgehen, werde ich im Rahmen der vorzustellenden Thesen noch eingehen.

---

4 Das Institut wurde zwar schon Anfang 1988 gegründet *(Teckenberg* 1988, S. 94), doch erst allmählich konnte die Arbeit beginnen. Außerdem ist man noch teilweise im Gebäudeannex der Hotels "Haus der Touristen" am Rande Moskaus untergebracht. Erst 1990/91 erfolgt der Umzug in das endgültige Domizil im Zentrum Moskaus.

5 Bei einer Zusammenkunft mit Mitarbeitern des Sonderforschungsbereich 3 in Frankfurt konnte man davon einen Eindruck erleben: Auf ihre Verwunderung, wieso nicht alle Ergebnisse des sozio-ökonomischen Sonderforschungsbereichs in Politikberatung umgesetzt werden und ihre Frage, wie denn beispielsweise die Simulationsmodelle zu neuen Besteuerungsformen bisher implementiert worden seien, wollten die westlichen Kollegen nur gewunden Auskunft geben.

6 Zu dieser Gruppe gehört auch eine meiner Informanten, Frau N. *Sorkaja* (s. Fußnote 1).

- Wenn auch eher im Rahmen der Wirtschaftswissenschaften verankert, so verdient doch die 1988 erfolgte Neugründung des Instituts für "Sozial-ökonomische Probleme der Bevölkerung" Beachtung. Räumlich und inhaltlich wird es dem "Zentralen ökonomisch-mathematischen Institut" in Moskau benachbart bleiben. Die Direktorin, Frau Natalija ***Rimaschewskaja,*** hat sich bisher allerdings mit durchaus sozialwissenschaftlich höchst relevanten Studien zu Fragen der Armut, der Zusammensetzung und Verteilung der Erwerbs- und Haushaltseinkommen hervorgetan.

- Laut Beschluß des Zentralkomitees der KPdSU soll bei der Akademie für Gesellschaftswissenschaften beim ZK der KPdSU ein **"Zentrum für soziologische Forschungen"** eingerichtet werden. Dessen Leiter, Zhan T. ***Toschtschenko***, ebenfalls Vizepräsident der Sowjetischen Soziologischen Vereinigung, hat sich auch bereits durch kritische Bürokratieanalysen hervorgetan. Insbesondere Analysen auf dem Gebiete der politischen Soziologie

  sind Gegenstand der Forschung.[7] Angestrebt werden auch Eingriffe in die parteigebundene regionale Politikberatung, insofern als man die bisherigen soziologischen Abteilungen bei den Gebietskomitees der Partei nun nach Möglichkeit der Koordination durch dieses Institut unterstellen will. Dies bedeutet einerseits Zentralisierung, andererseits die Chance, landesweite Informationen über politische Einstellungen und Betätigungen einer wissenschaftlichen Forschung zuzuführen.

Wenn schon keine Neugründungen, so sind doch einige Aktivitäten auch weiterhin erwähnenswert, vor allem, weil sie z. T. jüngeren Soziologen eine Arbeitsmöglichkeit geboten haben. Dazu gehört vor allem:

- Eine Gruppe von Sozialwissenschaftlern in Kiew um Vladimir I. *Paniotto* und Valerij ***Chmel'ko.*** Diese Soziologen, deren Alter um die 40 Jahre liegt, geben, wiewohl in die gegenwärtigen heftigen politischen Kontroversen in der Ukraine involviert, bisher unregelmäßig eine neue Zeitschrift heraus: "Philosophische und soziologische Gedanken". Dieses Journal dürfte sich aber nicht unbedingt an ein sozialwissenschaftliches Fachpublikum wenden.

Damit wäre ich beim nächsten Punkt. Es gibt bisher nur eine einzige Fachzeitschrift: Die "Soziologische Forschungen", wenngleich diese seit Januar 1990 monatlich (!) erscheint. Die Soziologen verfassen zur Zeit lieber publizistische Artikel. Dies ist, einem vermeintlichen Gebot des historischen Augenblicks folgend, zwar verständlich, dient jedoch nicht unbedingt der wissenschaftlichen Professionalisierung und läßt uns mehr über die jeweilige politische Richtung der Autoren erfahren als über die sowjetische Gesellschaft. Aktive Forschungsar-

7 Siehe den Artikel des Direktors, ***Toschtschenko***, und seines Stellvertreters, V. E. ***Bojkov***. In: S. I. 1990 Nr. 9, S. 22-29.

beit über dieselbe gerät darüber häufig vorerst ins Hintertreffen, auch wenn die großen, überregionalen Zeitungen wie schon früher auch selber Umfragen durchführen oder ihre jeweilige Klientel bitten, zu bestimmten Themen Stellung zu nehmen, was diese im allgemeinen auch in großem Umfange bereit sind zu tun. Immerhin ist dies auch ein Indiz für die breitere Resonanz sozialwissenschaftlicher Fragestellungen in Kreisen der sowjetischen intellektuellen Öffentlichkeit im Vergleich zu westlichen Gesellschaften. Ferner zahlt sich die Tätigkeit für Zeitschriften für die Autoren materiell aus, was angesichts des allseits beklagten Mangels ein weiterer, weniger "moralischer" Grund für die publizistische Tätigkeit der sowjetischen Soziologen sein dürfte. 1990 erhob das Institut für Meinungsforschung in einer Allunionsumfrage, welches die "interessanteste" Zeitung 1989 war. Folgende **Zeitschriften** konnten sich profilieren (genannt werden hier nur diejenigen, in denen auch Soziologen publiziert haben):

Ogonek	21,1%
Novyj Mir (an und für sich eine Literaturfachzeitschrift)	4,3%
Smena	4,0%
Drushba Narodov	0,4%
Znamija	0,3%
die Frage ist nicht zu beantworten	28,5%

Unter den **Zeitungen** galten folgende als "interessant":

Argumenty i Fakty	33,2%
Komsomol'skaja Pravda	16,4%
Trud (die Gewerkschaftszeitung)	10,0%
Izvestija	6,3%
Literaturnaja Gazeta (ein Pedant etwa zur "Zeit"; der Herausgeber ist der bekannte "Politologe" Fjodor Burlatzkij)	4,1%
Pravda	2,5%
Sovetskaja Rossija	1,1%
Moskovskie Novosti (im Westen auch als: Moscow News)	1,6%
die Frage ist schwer zu beantworten	11,8%

Seitdem im August 1990 ein "Gesetz über die Presse" verabschiedet wurde, ist die Zensur eingeschränkt. Häufig scheitern aber mutige Unternehmungen an vermeintlichem "Papiermangel", so etwa auch im Falle der renommierten Zeitschrift "Novyj Mir". Folgende neue, wöchentlich erscheinende Journale dürften auch für publizistisch orientierte Soziologen interessant sein:

- Glasnost'
- Kommersant (eine Zeitschrift für Werbung und Konsumforschung, die ihre Informationen interessanterweise nicht von der staatlichen Nachrichtenagentur Tass, sondern von der unabhängigen Agentur "Postfaktum" erhält);
- Tschelovek (heute eher übersetzt mit "Der Bürger" unter der Leitung von Boris *Jüdin).*

Nun möchte ich systematisch zu drei Thesen Stellung nehmen, die z. T. für den Wissenschaftsbetrieb in (ehemals) sozialistischen Ländern insgesamt gelten. Sie erklären uns eventuell auch die erschwerenden Bedingungen bei der Umstellung des Wissenschaftsbetriebs in der ehemaligen DDR, kennzeichnen zumindest aber die Forschungssituation der Soziologie in der UdSSR Ende 1990. Dabei wird sich zeigen, daß sich in den Grundvorstellungen und in den institutionellen Mustern bisher wenig verändert hat.

1. Immer noch gibt es keine autonome "scientific Community".

Bei hohem Zentralisierungsgrad, der allerdings auch für Frankreich typisch ist, empfingen früher sowjetische Wissenschaftler gewisse staatliche "Lehen" für die Erfüllung von Regierungsaufträgen. Daneben war im Rahmen des allgemeinen Institutsetats, der je nach Größe des Instituts eher reichlich ausfiel, was das sowjetische Gesetz des "big is beautiful" förderte, auch eigene (Grundlagen-) Forschung möglich. Kurz nach dem Amtsantritt *Gorbatschows* hat man versucht, die Institute - ebenso wie einige Betriebe - auf die sogenannte "Wirtschaftliche Rechnungsführung" umzustellen, d. h. sie sind gehalten, ihre Mittel selbst zu erwirtschaften. Dadurch hat der Anteil an Auftragsarbeiten eher zugenommen. Insbesondere ist man natürlich an Aufträgen seitens westlicher Gazetten und Rundfunkanstalten interessiert, um an die begehrten Devisen zu kommen. Die dann in Frage kommende Forschung spiegelt ähnliche Probleme kommerzieller Auftragsforschung wie im Westen wider, nur oft mit schlechterer Qualität.

Führte früher schon das Institut für Soziologie Telefon-Blitzumfragen im Moskauer Raum für amerikanische Zeitungen durch, so ist man heute auch am Institut für Meinungsforschung an solchen "Westkontakten" stärker interessiert. Immerhin gibt es dort konkretere Vorstellungen, wie allunionsweite, repräsentative Umfragen gemacht werden können, und man verfügt dort über ein weites Netz von kleineren Unterabteilungen in den einzelnen Regionen, um solche Umfragen landesweit zu koordinieren. Solche groß angelegten Studien dürften auch weiterhin im Zuge der "Begleitforschung für die Perestrojka" zentral finanziert werden, selbst wenn die Gewerkschaften als finanzstarke Partner ausfallen. Daneben zeigen aber die Kooperation des Instituts mit SINUS in München und die Herausgabe des Institutsbulletins in deutscher Sprache (wenn auch in gekürzter Fassung, Leske Verlag: "Öffentliche Meinung in der UdSSR"), wohin man sich zu wenden gedenkt. Das Institut führte - ebenfalls in Koope-

ration mit SINUS - einen Auftrag der Bonner Regierung aus, bei dem 4000 Personen in der gesamten UdSSR nach ihren Ansichten über die Deutschen befragt wurden (FAZ vom 17. 11. 1990).

In gewisser Weise gibt es auch bei Regierungsaufträgen wieder eine (leichte) Zensur. So führte ***Gorbatschow*** erstmalig in der sowjetischen Geschichte ein Dekret ein, welches die Verunglimpfung von Regierungsmitgliedern unter Strafe stellt (F.-C. *Schroetter* in: FAZ vom 9. 10. 1990). Erste Skandale gab es, als die Zeitung Argumenty i Fakty in einer Umfrage nicht ***Gorbatschow,*** sondern den damals noch unter den lebenden weilenden A. *Sacharow* an der Spitze der beliebtesten Politiker ermittelte (Autor der Studie: *Kopeljusch).*

Allerdings vergeben, wie gesagt, auch andere gesellschaftliche Organe sowie die Landeszeitungen Aufträge. Adressat der allgemeinen soziologischen Forschung ist damit aber nicht ein unabhängiges Wissenschaftskollegium. Gegenstand der Untersuchungen ist nicht die Überprüfung kontroverser Hypothesen. Dies drückt sich auch darin aus, daß es kein unabhängiges wissenschaftliches Gutachtergremium gibt. Unabhängige Forschungsförderung, wie sie unter Beachtung von allgemeinen Qualitätsmaßstäben im Rahmen etwa der VW-Stiftung oder der deutschen Forschungsgemeinschaft betrieben wird, ist in der Sowjetunion praktisch unbekannt. Bezeichnenderweise ist die DFG Ansprechpartner für die sowjetische Akademie der Wissenschaften, die aber als Versammlung von Honorationen bei ihrer Rekrutierung von Schwankungen des jeweiligen politischen Tageskurses abhängig ist und erst jetzt langsam ihre dem traditionalen Sowjetsystem getreuen Mitglieder auswechselt.

2. Es fehlt eine Ausdifferenzierung der Soziologie, die wissenschaftliche Konkurrenz beleben könnte. Dies geht einher mit Problemen bei der Nachwuchsrekrutierung.

Ich erwähnte bereits die eher zunehmende Konzentration auf Moskau, die auch in ehemals relativ freizügigen Nischen Forschende in Leningrad und insbesondere in Novosibirsk dazu veranlaßt, nach Moskau zu ziehen.

Entscheidender scheint mir aber die sich je nach vorhandenem Budjet verstärkende Tendenz zu Großinstituten zu sein. Die Neigung, eine für kleinere Institute günstigere gemeinsame Infrastruktur aufzubauen, ist dagegen gering. Ähnlich wie in der DDR kommt das auch dadurch zustande, daß auch der Mittelbau der Institute unkündbar ist, während Direktoren und die umfangreichen Vizedirektorenstellen je nach politischer Wetterlage durch Wahl gewechselt werden können.[8]

---

8 Sie erhalten aber in der Regel eine Apanage in einem weniger "zentralen" Institut.

Wo sollten gekündigte oder junge, engagierte Nachwuchswissenschaftler aber auch unterkommen? Die vielen kleineren und wenig koordinierten Sozialforschungsstellen bei den industriellen Großbetrieben sind - besonders wenn sie nicht im Moskauer Raum liegen - nicht sehr attraktiv, die Soziologen dort beim Management nicht unbedingt beliebt. Viele sozialwissenschaftlich Interessierte der mittleren Generation wanderten unter *Brezhnev* in die Politik oder zu Zeitungen ab. So kommt es, daß die heute wieder in Spitzenstellungen akkreditierten Frauen und Männer der Geburtsstunde der sowjetischen Soziologie aus der Ära nach ***Chruschtschow*** sind und etwa ein Alter haben, wie der heutige Regierungschef selbst.[9] Besonders die mittlere Generation fehlt heute oder kommt im System der vielfältigen, direkten Beziehungen zu den Macht und Geld habenden gesellschaftlichen Organisationen (noch) nicht zum Zuge. Viele sind auch während des im Zuge der Liberalisierung verstärkt anhaltenden "Brain Drains" ausgewandert.

Für kleine Forschungsinstitute besteht auch aufgrund der eigenartigen Vorstellung oder Nichtvorstellbarkeit von koordinierenden Infrastruktureinrichtungen kaum eine Chance. Die Großforschungsinstitute sind nicht vernetzt. Selbst in Moskau doublieren sie ihre Standardausrüstung ohne Anschluß an zentrale Großrechner und schreiben sogar ihre Programme jeweils selbst. SPSSx gibt es im Saslawskaja-Institut erst seit Frühjahr 1990, und die Bereitschaft zum Programm- und Erfahrungsaustausch zwischen den Instituten ist nicht sehr ausgeprägt. Auch ein zentrales Datenarchiv gibt es nicht - begründet wird dies neuerdings - wie sollte es auch anders sein - mit "Datenschutz".

Die mangelnde Eigenkoordination der Soziologie ist z. T. ein Ergebnis jahrelanger Plandirektiven "von oben". Doch auch heute kommen als koordinierende Organe vor allem die Fachverwaltungen in Frage, zumindest solange ihnen zugetraut wurde, verbindliche Richtlinien für die Erfassung der grundlegenden Statistik festzulegen. Nachdem aber die Arbeit des Staatskomitees für Statistik kritisiert worden ist und die Unzulänglichkeit der im allgemeinen alle zehn Jahre stattfindenden Volkszählung (zuletzt 1979) gerügt wurde, hat man heutzutage weniger brauchbare Aggregatdaten als je zuvor. Die jetzt erst teilweise erschienenen Bände der Volkszählung von 1979 weisen im Gegensatz zu früheren Publikationen keine Angaben über den Beruf aus. Wegen mangelnder Vorbereitung und Fehlern bei der Durchführung setzt man keine großen Hoffnungen in die Volkszählung von 1989. Streit zwischen verschiedenen Richtungen und Instituten verhindert sowohl die Einigung auf bestimmte gemeinsame Standards, so etwa auf eine verbindliche, allgemeine Berufsklassifikation. Ein gewisser Minimalkonsens, etwa über eine Standarddemographie (einschließlich der Berufsvercodung), wäre allerdings Voraussetzung für gemeinsame Infrastruktureinrichtungen. Die ungenügende oder

---

9 Siehe dazu die von mir verfaßten Biographien von 44 sowjetischen Soziologen der Gegenwart. In: W. ***Bernsdorf*** (Hg.) (1984), Internationales Soziologenlexikon. Band 2. Neue Ausgabe. Berlin, Stuttgart: Enke Verlag.

kaum vergleichbare Standarddemographie, die eine Basis für die Bildung unabhängiger Variablen bilden könnte, verführt dazu, lediglich zu beschreiben. Erklärungen des Verhaltens oder der Diversifikation der Meinungsbildung sind dagegen selten. Ähnliches Desinteresse an übergreifenden Strukturbeschreibungen der Gesellschaft gilt natürlich auch für westliche Auftragsarbeiten von Marktforschungsinstituten, die mehr auf die "Meinungen über die Welt und die Welt der Meinungen" (so der Titel eines Buchs von Boris ***Gruschin)*** achten, als daß sie an einer sozialstrukturellen Erklärung der Einstellungen und Dispositionen interessiert wären.

Typisch für die mangelnde Ausdifferenzierung ist, daß es bisher nur eine einzige soziologische Fachzeitschrift gibt, die redaktionell vom Akademieinstitut für Soziologie betreut wird. Die Bemühungen sowjetischer Kollegen, sich in internationalen Fachzeitschriften zu artikulieren, sind - anders als bei ihren mitteleuropäischen Nachbarn - bisher wenig ausgeprägt. So gibt es dann sehr wohl einen Wettbewerb um Gelder, besonders nun auch aus dem Ausland, jedoch wenig Konkurrenz um wissenschaftliche Reputation. Dies heißt nicht, daß im Zuge der Neuorientierung gerade jene Soziologen in führende Stellen (wieder) aufgerückt sind, die schon zu Anfang der 70er Jahre unter den gegebenen Bedingungen gute empirische Studien vorgelegt haben. Jedoch werden die "Standards of Excellence" eher durch moralisch-ideelle Kriterien gesetzt oder durch das jeweilige politische Gebot der Stunde.

Die Sowjetische Soziologische Vereinigung, der Berufsverband der Soziologen mit T. *Saslawskaja* an der Spitze[10], kann zwar als Pendant zur Deutschen Soziologischen Vereinigung gesehen werden, funktioniert jedoch nach anderen Regeln. So gibt es z. B. keine Soziologentage, die auch jüngeren Kollegen eine Plattform für eine beachtenswerte "Kür" im Rahmen der Profession bieten könnten. Eine der Funktionen der Vereinigung ist die Außendarstellung der sowjetischen Soziologie, insbesondere auf den Weltkongressen für Soziologie. Die Rangordnung innerhalb des Verbands bestimmt im wesentlichen, wer zu den jeweiligen "Reisekadern" zählt. Die auf internationalen Tagungen augenscheinliche Geschlossenheit der Vereinigung, die größer ist als etwa bei westlichen Soziologieverbänden, zerfällt, wenn einzelne Kollegen je nach der Verfügbarkeit von entsprechenden Beziehungen und Einladungen ins westliche Ausland reisen. Bekanntheitsgrad und Beziehungen sorgen dafür, daß heute (noch) vor allem ältere Kollegen zu den "Reisekadern" zählen, und dies, obwohl die Fremdsprachenkompetenz bei jüngeren Soziologen in vielen Fällen besser wäre.

---

10 Als Vizepräsidenten wurden gewählt (alle Professoren der philosophischen Fakultät, s.u.): Vilen N. ***Ivanov*** (im Oktober 1988 noch Direktor des Instituts für soziologische Forschungen), Vladimir A. ***Jadov*** (Moskau, Direktor des Instituts für Soziologie der Akademie), Boris V. ***Rakitskij*** (Moskau), Ovsej I. ***Schkaratan*** (Moskau, Institut für Ethnographie der Akademie), Mikk Ch. ***Titma*** (Tallinn), Shan T. ***Toschtschenko*** (Moskau). G. V. ***Osipov***, der immerhin korrespondierendes Mitglied der Akademie der Wissenschaften ist und damit eine hohe Reputation nach sowjetischem Verständnis innehat und auch im Westen gut bekannt ist, gehört nicht mehr zum engeren Präsidium.

Bei aller Bewunderung, die sowjetische Intellektuelle besonders dem amerikanischen Wissenschaftssystem entgegenbringen, wird kaum gesehen, daß dieses nach ganz anderen Kriterien funktioniert und daß sich dort heute eher Soziologen aus dem asiatischen Raum aufgrund ihres Leistungsstrebens durchsetzen als Forscher aus der Sowjetunion, denen ein auf Marktprinzipien beruhendes Wissenschafts System eher fremd ist.
Dies bringt mich zur nachten These:

3. Statt Marktkonkurrenz "Standespolitik" und politische Patronage.

Wie Max Weber schon feststellte, entstehen Ständegesellschaften eher dann, wenn der Markt zurückgedrängt ist.

Zwecks Erlangung von Forschungsgeldern spielen weniger die wissenschaftliche Qualifikation, sondern vielmehr Beziehungen zu öffentlichen Gremien und deren Mitarbeiter eine Rolle. Nicht, als ob es das in westlichen Gesellschaften nicht gäbe - jedoch im System der patrimonialen Protektion von Wissenschaft durch verschiedene gesellschaftliche Organisationen nicht nur durch die Partei, sondern auch durch Gewerkschaften etc. ist der direkte Zugang zu den entsprechenden "Flügeln" der Organisationen entscheidend. Je nach politischem Klima kann dann auch die Protektion wieder entzogen werden, wie man an der nun fehlenden Bereitschaft der eher sowjet-traditional orientierten Gewerkschaften sieht, fürderhin das reformbewußte Saslavskaja-Institut zu unterstützen.

Zweifellos ist die Rolle der Soziologie bei der Politikberatung in der UdSSR größer als etwa in der Bundesrepublik. Vor allem das Institut von *Saslawskaja* betreibt mit den landesweiten Meinungsumfragen direkte Begleitforschung zur Perestrojka. Der Besuch von Frau *Saslawskaja* bei Frau *Noelle-Neumann* bekundet ihr Interesse an einer Soziologie, die weitgehend durch Regierungs- oder Medienaufträge finanziert wird. Die Bemühungen um politische Patronage hängen natürlich mit der mangelnden Unabhängigkeit von Wissenschaft in der UdSSR zusammen und auch andere Institutionen erhalten Regierungsmittel. Die Gründung des Soziologischen Zentrums unter der Leitung von *Toschtschenko* (s. o.) bei der Partei beweist, daß auch letztere von der Bedeutung der Soziologie überzeugt ist. Vielfältige persönliche Beziehungen vermitteln Soziologen mit unterschiedlichen Einflußkanälen in die Politik.[11]

---

11 Wie auch immer die Tatsache zu werten ist, daß Raissa ***Gorbatschowa*** bei G. V. ***Osipov*** eine Arbeit geschrieben hat, es hat letzteren nicht vor einer gewissen Degradierung bewahrt. Zweifellos kann man aus dieser publizistisch aufgewerteten Beziehung nicht die verstärkte Förderung der Soziologie erklären. Da ***Gorbatschow*** selbst Jurist ist, müßte die Jurisprudenz eine ähnliche Ausweitung erfahren haben.

Ein Beispiel standesgemäßer Profilierung der Intelligentsija ist die publizistische Tätigkeit der Soziologen, die je nach politischer Couleur in dem sich erfreulicherweise erweiternden Spektrum der Zeitungen ein bestimmtes Publikum erreichen und dafür dann bezahlt werden. Die Orientierung der Soziologen und ihr involviertes Auftreten in der Öffentlichkeit ist teilweise den von A. **O. *Hirschman*** beschriebenen Zyklen von "Engagement und Enttäuschung" unterworfen.

Bei den Stellungnahmen spielen neben der politischen Richtung auch offenbar wissenschaftsfremde Kriterien eine Rolle, wie die Affiliation mit eher vorsozialistischen Spannungslinien, wie der zwischen "Slawophilen" und "Westlern". Zu diesen Orientierungen zählen auch die heute besonders deutlich werdenden Schattierungen bestimmter, mehr oder weniger chauvinistischer, Nationalismen.

In Abwesenheit einer Marktkonkurrenz, die auf wissenschaftlicher Qualifikation beruht und bei mangelnder Einbindung der sowjetischen Soziologie in die entsprechenden westlichen Publikationsorgane wird innerhalb der als solche einheitlich auftretenden Sowjetischen Soziologischen Vereinigung - wenn es um zentrale Forschungsgelder geht - durchaus nach funktionalen Äquivalenten zur Aufrechterhaltung bestimmter Standards gesucht. Diese werden aber spezifisch sowjetisch durchdekliniert:

Wenn Stände nach Max *Weber* auch durch die Anmutung einer bestimmten Lebensführung gekennzeichnet werden können, so sind die Verabschiedung eines "Professionellen Kodex des Soziologen" und die Einsetzung eines "Rats für professionelle Ethik" durch die soziologische Vereinigung aufschlußreich für die Standespolitik. Wenngleich es einen festgeschriebenen "Code of Ethics" auch neuerdings in der amerikanischen Soziologenvereinigung gibt, ist dessen Funktion in der UdSSR doch grundlegender. Im Kodex (S. I. 1988, Nr.5, S.95-104) werden die Soziologen immer noch aufgefordert, sich im Sinne der marxistisch-leninistischen Soziologie der Parteilichkeit zu befleißigen. Da sie aber auch gebeten werden, "forschend nach der Wahrheit zu suchen" und ihre Meinung unabhängig von Konjunkturen und Autoritäten zu äußern, wird hier "Parteilichkeit" im Sinne einer großen "catch-all" Volkspartei verstanden, eine Auffassung, die durch den Abdruck einer Diskussion über den "sozialistischen Pluralismus" in derselben Nummer der Fachzeitschrift noch unterstrichen wird. Der Rat für Fragen der professionellen Ethik, der aus 15 Personen, auch aus früher verfemten Soziologen (z. B. Jurij A. *Levada),* besteht, soll darüber wachen, daß die wissenschaftlichen Standards eingehalten werden. Die Nichtbeachtung des Kodex kann zum Ausschluß aus der sowjetischen soziologischen Vereinigung führen, was praktisch einem Publikationsverbot - oder analog zu früher ähnlichen Praktiken im Schriftstellerverband - einem Berufsverbot gleichkommt, zumal es keine ausgeprägte Fachzeitschriftenlandschaft gibt. In dem anschließenden Aufsatz zu diesem "Moralkodex", für den auch V A. *Jadov* verantwortlich

ist, wird deutlich, daß man nach Regelungen sucht, die die Selbststeuerung und Autonomie des Faches garantieren und es vor dirigistischen und administrativen Übergriffen schützen sollen. Allgemein und international akzeptierte Standards der Wissenschaftlichkeit müssen dann erst für jede Sparte einzeln eingeführt werden, wenn das betreffende Fach früher starken administrativen Vereinnahmungsbestrebungen ausgesetzt war.

Insgesamt hat man den Eindruck, daß zunächst die ständische Rangordnung der Wissenschaftssparte "Soziologie" festgelegt werden muß. Der entsprechende "Status" schafft dann gewisse Privilegien und eine weitgehende Absicherung der Beschäftigungsverhältnisse und der Aufgabenfelder. Berufsständische Forderungen stehen zur Zeit noch im Vordergrund, und deren Durchsetzung benötigt im Augenblick noch Zeitaufwendungen, die zunächst nicht für Forschungsarbeiten zur Verfügung stehen. Vielleicht kann man aber die Bemühung um etatistische Festschreibung verstehen, nachdem die Jahre 1975 bis 1985 der Soziologie zunächst einen Rückschritt beschert hatten. Heftige Debatten über die Wahl der Direktoren und deren Stellvertreter bestimmen häufig den Tagesablauf in den Instituten. Daneben beherrschen - in der gegenwärtigen Situation verständlich - politische Disputationen die Tagesordnung. Dies kontrastiert mit westlichen Effizienzvorstellungen, nach denen sich eine junge Wissenschaft oder ihre Anhänger "marktgerecht" einbringen müssen und in Konkurrenz zu anderen durchsetzen sollen.

4. Kontakte mit dem Westen. Ein Ausblick.

Marktwirtschaftliche Aspekte, aber auch bestimmte Beziehungen bestimmen heute vor allem die Konkurrenz um Anteile am westlichen Medienmarkt, weniger aber sowjetische Bemühungen, im Rahmen westlicher sozialwissenschaftlicher Publikationen Fuß zu fassen. Angesichts der gegenwärtigen Umbruchsituation in der UdSSR ist die vorherrschende Binnenorientierung der sowjetischen Soziologie zum Teil verständlich. Auch die Westorientierung des Saslawskaja-Instituts richtet sich vornehmlich an einer Meinungsforschung für westliche Medien aus. Die Westkontakte der sowjetischen Grundlagenforschung, die ja als Institution durchaus vorhanden ist, nehmen sich dagegen eher spärlich aus oder beschränken sich auf Reiseaktivitäten prominenter Soziologen. Insbesondere mit den Amerikanern bestehen allerdings auch Pläne für international vergleichende Sozialforschungsprojekte. Die Kooperation wird über den Atlantik auch durch eine Institution erleichtert, nämlich durch das International Russian Exchange Program (IREX, Verantwortlicher: Wesley A. *Fisher*, New York). Im Verhältnis zur Bundesrepublik bestehen sporadische, bilaterale Austauschbeziehungen zwischen Instituten und immerhin ist es dem Zentralarchiv nun gelungen, das Datenband einer Meinungsumfrage des Saslawskaja-Instituts zu erhalten. Dies wäre noch letztes Jahr fast undenkbar gewesen. Jedoch gibt es nun neue Probleme, die um den Bereich des Datenschutzes kreisen; auch die Weitergabe von Daten, die bereits von anderen Stellen - etwa den kommerziellen

Instituten - genutzt wurden, dürfte zunehmend problematisch werden. Im Bereich der allgemeinen Soziologie, die im Vergleich zu kommerziellen Instituten ein weniger finanzstarker westlicher Partner ist, erweisen sich für die Kooperation die unverändert kaum überwindbaren Barrieren der internationalen Kommunikation, auch auf der einfachsten Ebene des Brief - und Fernmeldewesens, als große Hemmschwelle. Seitens der Bundesrepublik erschweren stark gestiegene Kosten eine Reise in die UdSSR, bzw. teilweise lange Antragszeiten bei einer Bewerbung über das bestehende Austauschprogramm der Deutschen Forschungsgemeinschaft mit der Akademie der Wissenschaften in der UdSSR, den Meinungsaustausch vor Ort. Seit diesem Jahr besteht prinzipiell die Möglichkeit, daß auch sowjetische Kollegen auf private Initiative Bundesbürger einladen. In diesem Falle sind umfangreiche behördliche Abwicklungen notwendig, die ebenfalls durch Kommunikationshemmnisse erschwert werden. Deshalb ist die Absicht des Bundesforschungsministeriums (BFMT) sehr zu begrüßen, in Moskau eine Art Koordinationszentrum für den Bereich der Wirtschafts- und Sozialwissen-Schäften einzurichten.[12] Dies würde teilweise die eminenten Kommunikationsprobleme eliminieren und könnte dazu beitragen, daß auch jüngeren Kollegen aus West und Ost die Möglichkeit der Kontaktaufnahme geboten wird. Vorschläge zur Kooperation, etwa zur Herausgabe einer gemeinsamen Fachzeitschrift, sei es in Englisch oder in Deutsch, (seitens Herrn *Schkaratans),* gibt es jedenfalls zur Genüge.

Wolfgang Teckenberg
Institut für Soziologie der Universität Heidelberg,
Sandgasse 9
6900 Heidelberg 1

12 Siehe dazu: Heinrich ***Vogel***, "Deutsches Wissenschaftszentrum Moskau" (Entwurf eines Memorandums). Köln, Bundesinstitut für ostwissenschaftliche und internationale Studien. September 1990.

Conference on Correspondence Analysis

Call for papers

The Zentralarchiv für Empirische Sozialforschung is organizing a Conference on correspondence analysis with emphasis on the multiple case and related methods. The Conference is to be held on

May 21-24,1991, at the University of Cologne, Germany.

The Conference is organized in collaboration with Professor Walter *Kristof,* Institut für Soziologie, University of Hamburg. We are happy to announce that Professor Michael **J. *Greenacre*** has agreed to present a detailed introduction and to discuss the most recent developments. Time will also be allocated for a demonstration of Computer Software used in correspondence analysis.

Correspondence analysis is a multivariate Statistical technique which analyses categorical data, typically cross-tabular data from social surveys, and attempts to display such data graphically to assist interpretation and understanding of the data. In his introduction Michael ***Greenacre*** presents a linked series of papers on the geometric interpretation of correspondence analysis and recent developments in multiple correspondence analysis. These papers will cover such topics as the geometric concepts in correspondence analysis, the use of supplementary points, the interpretation of row and column contributions, the application of correspondence analysis to multiway data and the complementary use of correspondence analysis and Cluster analysis.

This introduction will cover approximately the first three days. We wish to invite papers by others to be presented on the final day of the meeting, especially on newer developments, interesting applications, and/or controversial issues. Please send abstracts to Jörg *Blasius,* Zentralarchiv für Empirische Sozialforschung (EARN: BLASIUS at DKOZAl).

The deadline for contribution is February 15,1991; for participation only is April 15,1991. The Conference fee is DM 50,-.

Participants who intend to present applications are invited to bring their data sets with them on Standard floppy disks.

Empirische Sozialforschung über Entwicklungsländer (ESE): Datensammlung und vergleichende Datenanalyse

Ankündigung einer Tagung im Zentralarchiv

Das Forschungs- und Datensammlungsprojekt 'Empirische Sozialforschung über Entwicklungsländer' (ESE), das gemeinsam von der Arbeitsstelle für Entwicklungsländerforschung und dem Zentralarchiv getragen wird, plant für das kommende Jahr eine Fachtagung, auf der wir Projektergebnisse vorstellen und Methodenprobleme der empirischen Entwicklungsländerforschung mit interessierten Kolleginnen und Kollegen diskutieren möchten.

Die Tagung soll am 20./21. Juni 1991 in Köln stattfinden.

Beiträge sind u.a. zu folgenden Themenbereichen vorgesehen:

- Probleme quantitativ-empirischer Erhebungen in Entwicklungsländern
- Datenqualität von Entwicklungsländerstudien - Erfahrungen bei der Archivierung von Datensätzen
- Möglichkeiten und Grenzen vergleichender Sekundäranalysen von informal sector-Studien
- Standardisierung von Erhebungsinstrumenten und Fragebogenmodulen als Mittel zur Verbesserung von Datenqualität und Vergleichbarkeit von quantitativen Studien
- Fortbildung in Erhebungstechniken und Datenanalyse am PC - eine praxisnahe Möglichkeit der Wissenschaftskooperation mit Entwicklungsländern

Interessenten sind eingeladen, sich schon jetzt - auch mit Vorschlägen für eigene Beiträge - zu wenden an:

Dr. Christoph Reichert
Arbeitsstelle für Entwicklungsländerforschung (AEF)
Richard Wagner-Str. 39
5000 Köln 1
Tel. 0221/4704078

oder

Dietmar Kaross
Zentralarchiv für empirische Sozialforschung (ZA)
Bachemer Str. 40
5000 Köln 41
Tel. 0221/4769451

Über die Arbeit des Projekts 'Empirische Sozialforschung über Entwicklungsländer' informiert auch der bisher zweimal erschienene ESE Newsletter, der auf Wunsch zugeschickt wird.

Bücherhinweise

Jörg Blasius, Jens S. Dangschat
Gentrification: Die Aufwertung innenstadtnaher Wohnviertel
Frankfurt a.M.: Campus 1990
423 Seiten

In einer Zeit knappen Wohnraumes und insbesondere des Mangels an preiswertem Wohnraum in Großstädten kommt der Aufwertung innenstadtnaher Wohngebiete ("Gentrification") besondere Aufmerksamkeit zu. Die Gründe der Aufwertung liegen sowohl in einer veränderten Nachfrage als auch in Verschiebungen auf der Angebotsseite.

In dem Buch von *Blasius und Dangschat* werden globale wirtschaftliche und gesellschaftliche Entwicklungen, die nationale Steuerpolitik und veränderte demographische Bedingungen als Ursachen für das Aufkommen der Nachfrage nach attraktivem innenstadtnahem Wohnraum beschrieben. Auch die Angebotsseite wird durch Steuer- und Wöhnungspolitik auf nationaler und kommunaler Ebene sowie durch lokale Planung bestimmt, die eine Reinvestition in den Altbaubestand attraktiv machen.

Die Folgen sind einerseits der Erhalt des innenstadtnahen Wohnungsbestandes, andererseits die Verdrängung alteingesessener Bewohner mit den entsprechenden sozialen Auswirkungen. Die Baustruktur wird modernisiert und instandgesetzt, alte Bestände werden abgerissen und durch Neubauten ersetzt, Baulücken werden geschlossen, Dächer werden ausgebaut, das Wohnumfeld wird verbessert. Die baulichen Umstrukturierungen führen zu einem deutlichen Anstieg der Mieten, einer Umwandlung von Miet- in Eigentumswohnungen sowie zu einer Aufwertung der Infrastruktur.

Die sozialen Folgen sind eine zunehmende Segregation der Wohnstandorte, die Verdrängung unterschiedlicher einkommensschwacher Gruppen an Wohnorte mit schlechter Standortgunst, der Verlust der vertrauten Wohnumgebung mit den entsprechenden psychosozialen Folgen für die Betroffenen. Dieser Prozeß wird daher auch häufig als "qualitativer Bevölkerungsaustausch" bezeichnet. Gentrification wird damit zur räumlichen Widerspiegelung und

zum Symbol wachsender gesellschaftlicher Unterschiede und provoziert zunehmende soziale Konflikte.

Ferner wird in diesem Band über die vorhandenen und die notwendigen planerischen und kommunalpolitischen Instrumente sowie ihre Anwendung diskutiert, um die Aufwertung innenstadtnaher Wohngebiete abzubremsen und zu kontrollieren. Solange jedoch die Aufwertung innenstadtnaher Wohngebiete von Politikern und Planern als erfolgreiche Gegenstrategie gegenüber einer Suburbanisierung gesehen und dementsprechend forciert wird, werden die vorhandenen Instrumente nur unvollständig und zögerlich eingesetzt.

Ausgangspunkt dieses Bandes war eine Tagung am Zentralarchiv für empirische Sozialforschung in Köln vom 12. -14.4.1989, die die beiden Herausgeber vorbereiteten. Ziel dieser Tagung war, Wissenschaftler aus Forschungseinrichtungen sowie Planer und Praktiker aus unterschiedlichen bundesdeutschen Großstädten an einen Tisch zu bringen. Ein Teil der Vorträge wurde für diese Veröffentlichung überarbeitet. Außerdem wurde eine Reihe weiterer Personen aufgefordert, ergänzende Aufsätze zu verfassen. Es handelt sich in allen Fällen um Originalbeiträge.

Dieses Buch, das im Campus Verlag Frankfurt erschienen ist, ist auch im Zentralarchiv erhältlich. Benutzern unseres Hauses können wir dieses Buch zum **Sonderpreis von DM 35,-** (Ladenpreis DM 78,-) anbieten. Richten Sie schriftliche Bestellungen bzw. Bestellungen via BITNET (BLASIUS @ DK0ZA1) an das Zentralarchiv, z.Hd. Herrn Jörg *Blasius*. Die Lieferung erfolgt gegen Rechnung.

Aus dem Inhalt:

Die Aufwertung innenstadtnaher Wohngebiete - Grundlagen und Folgen
Jens S. Dangschat, Jörg Blasius

Der Einfluß von ökonomischen und sozialen Prozessen auf die Gentrification
Hartmut Häußermann

Die Gefährdung preisgünstigen Wohnraums durch wohnungspolitische Rahmenbedingungen
Volker Kreibich

Geld ist nicht (mehr) alles - Gentrification als räumliche Segregierung nach horizontalen Ungleichheiten
Jens S. Dangschat

Stadterneuerung und Gentrification am Beispiel Frankfurts
Werner Wingenfeld

Der Verlust preiswerten Wohnraums - kann er abgeschwächt, können seine Folgen abgemildert werden? Das Beispiel München
Günter Distler

"Friede den Hütten ...?" Traditionelle und erweiterte Handlungsansätze zur Sicherung preiswerter Wohnverhältnisse
Rolf Froessler, Klaus Seile

Die Erhaltungssatzung als Mittel zum Erhalt der Sozialstruktur
Ursula Stein

Gentrification - Bewertung aus planerischer Sicht
Walter Richter

Die Aufwertung innenstadtnaher Wohnviertel in Köln
Arndt Schulz

Gentrification: Stadtteile im Konflikt - Beispiele aus Hamburg
Carl-Heinrich Busse

Gentrification und Stadterneuerung - Dargestellt am Fallbeispiel Stuttgart
Wahlburg Ratz-Hoberg

Zur Messung von Gentrification - Anmerkungen zu einem komplexen Forschungsdesign
Harald Rohlinger

Sozialwissenschaftliche Mehrebenenmodelle und Gentrification
Johannes Huinink

"Das gute Wohngefühl ist weg!" - Wahrnehmungen, Bewertungen und Reaktionen von Bewohnern im Vorfeld der Verdrängung
Monika Alisch, Wolfgang zum Felde

Soziale und strukturelle Folgen von Modernisierungen in innenstadtnahen Gebieten
Reinhard Wießner

Bestimmungsfaktoren, Wirkungszusammenhänge und Folgen der Umwandlung von Miet- in Eigentumswohnungen
Winfried Killisch, Reinhold Gütter, Michael Ruf

Gentrification und Lebensstile
Jörg Blasius

Junge Erwachsene auf großstädtischen Wohnungsmärkten - Zugänge, Plazierungsprobleme, Ausgrenzungen
Wolfgang Gaiser, Hans-Ulrich Müller

Wohnungsneubau in der Mannheimer Innenstadt
Stefan Krämer, Katrin Zapf

Dieter Jauf mann, Ernst Kistler, Günter Jänsch
Jugend und Technik: Wandel der Einstellungen im internationalen Vergleich.
Frankfurt a. M.: Campus 1989

Im Geleitwort zu diesem Buch schreibt der Bundesminister für Forschung und Technologie Dr. Heinz ***Riesenhuber*** u.a.:
Veränderungen der öffentlichen Meinung in der Einschätzung von Chancen und Risiken des technischen Fortschritts werden seit einer Reihe von Jahren lebhaft diskutiert und in zunehmendem Maße als wichtige Indikatoren für gesellschaftliche Grundeinstellungen bewertet. Da sich die Erkenntnis durchgesetzt hat, daß der technische Fortschritt nicht nur ein natur- und ingenieurwissenschaftlicher, sondern mindestens in gleichem Maße ein gesellschaftlicher Prozeß ist, wird die Einstellung zur Entwicklung und Nutzung von Technik zu einem ganz entscheidenden Bestandteil der Tendenzen unserer von Wissenschaft und Technik geprägten Zivilisation.

Meinungs- und Sozialforschung haben die damit zusammenhängenden Fragen schon verhältnismäßig früh in der Bundesrepublik aufgegriffen. Das vielfältige und reiche demoskopische Originalmaterial war aber bisher noch keiner systematischen, längerfristig aufbauenden Untersuchung unterzogen worden: Vielfach wurde bezweifelt, daß bei der notwendigerweise sehr verschiedenartigen Ausgangslage des Umfragematerials - nach zeitlichem, methodischem und thematischem Umfeld - eine derartige Sekundäranalyse überhaupt mögüch wäre. In noch größerem Maße galten diese Zweifel dem internationalen Vergleich, der interkulturellen Vergleichbarkeit von demoskopischen Ergebnissen zum Thema Technikakzeptanz.

Mit dem Buch stellen die Autoren den Versuch vor, diese Forschungslücke zu füllen. In ständigem Kontakt mit der Praxis der Meinungsforschung und der Methodenentwicklung in der Sozialforschung, u.a. in Form von drei internationalen Forschungskolloquien, wurde erstmals in diesem Umfang das verfügbar gemachte Material der deutschen Meinungsforschung und zahlreiche ausländische Umfragen unter besonderer Berücksichtigung der jüngeren Generationen auf Veränderungen der öffentlichen Technikakzeptanz untersucht. Es wurde ein Pioniervorhaben durchgeführt, das einen wichtigen Beitrag zum methodischen und inhaltlichen Fortschritt nicht nur auf dem Gebiet der sozialwissenschaftlichen Sekundäranalyse darstellt, sondern auch durch Primärerhebungen ergänzt und weitergeführt wurde.

Empirische Sozialforschung 1989

Eine Dokumentation von
Steffen Kühnel und Harald Rohlinger

Frankfurt a. M.: Campus, 1990

Die soeben erschienene Dokumentation gibt einen Überblick über sozialwissenschaftliche Forschungsprojekte, die Methoden der quantitativen empirischen Sozialforschung anwenden. Dargestellt werden geplante, laufende und 1989 abgeschlossene Projekte. Ziel der Dokumentation ist es, einen frühzeitigen Informationsaustausch zwischen den Forschern zu ermöglichen. Neben Anregung zur Kooperation kann die Dokumentation auch genutzt werden, um bei der Planung eigener Projekte Doppelarbeit zu vermeiden. Schließlich können die aufgeführten Projektberichte auch Anregungen zu Sekundäranalysen bieten.

Grundlage der Dokumentation ist eine Erhebung, die das Informationszentrum Sozialwissenschaften (Bonn) von Oktober bis Dezember 1989 durchgeführt hat. Dazu wurden Institutionen in der Bundesrepublik und der deutschsprachigen Schweiz angeschrieben. Aus Österreich liegen Projektberichte vor, die die Sozialwissenschaftliche Informationsstelle in Wien erhoben hat. Die Erhebung ist die umfassendste dieser Art im deutschsprachigen Raum.

Der Rücklauf aus den angeschriebenen Institutionen entspricht im wesentlichen den Ergebnissen der Vorjahre. Gut ein Drittel (34 %) der Institutionen meldet eigene Forschungsarbeiten. Da in der Regel mehrere Projekte gemeldet werden, übersteigt die Zahl der gemeldeten Projekte die Zahl der Institutionen. Insgesamt enthält die Dokumentation 1989 1.084 Projektbeschreibungen.

Einige Ergebnisse der diesjährigen Erhebung

Die Analyse der Struktur der gemeldeten Projekte zeigt ein weitgehend ähnliches Bild wie in den vergangenen Jahren. Bei der Klassifikation der Projekte nach der Art der **Institution,** an der ein Forschungsprojekt durchgeführt wird, dominiert auch 1989 der Hochschulbereich.

637 (= 64 %) von 996 einordbaren quantitativen Projekten werden an einer Hochschule durchgeführt.

Die akademische Ausrichtung der Sozialforschung kann nicht nur am Anteil der Hochschulforschung abgelesen werden. Sie zeigt sich auch daran, daß 62 % der Projektleiter einen Professorentitel aufweisen.

Die gemeldeten Forschungsprojekte dienen in erheblichem Maße der wissenschaftlichen Qualifikation. So wird bei der Frage nach der **Art der Forschungsprojekte** von 21% der hierzu vorliegenden Antworten aus 996 quantitativen Projekten berichtet, daß das Projekt (zumindestens auch) der Erreichung eines akademischen Grades dient.

Seit geraumer Zeit ist die Verwirklichung der Chancengleichheit von Frauen und Männern ein stark diskutiertes Thema in der Öffentlichkeit. Das Augenmerk wird dabei auch auf den Anteil der Frauen in den verschiedenen Berufen gerichtet, bilden Berufe doch eine der wichtigsten Ressourcen für die Erlangung höherer sozialer Positionen in der Gesellschaft. Dabei lautet durchweg der Befund: Frauen sind in einer Reihe von Berufen entweder gar nicht vertreten oder nur zu einem sehr geringen Anteil. So sind insbesondere Stellungen in leitenden Positionen für Frauen offensichtlich nur schwer zugänglich. Warum das so ist und wie die Chancen der Frauen zu verbessern sind, darüber wird in der öffentlichen Erörterung heftig debattiert.

Die für die "Empirische Sozialforschung 1989" gemeldeten Projekte erlauben es, erste Ansätze zu einer Bestandsaufnahme der **Beteiligung von Frauen und Männern** an den Forschungsarbeiten zu liefern. So ergibt eine Aufteilung dieser Personen nach dem Geschlecht, daß nur 10 % der Leiter und Betreuer eines Projektes Frauen sind. Bei den Projektbearbeitern ist der Anteil der weiblichen Mitarbeiter mit 30 % zwar deutlich höher, entspricht aber nicht dem Anteil der Frauen an der erwerbstätigen Bevölkerung insgesamt, der in der Bundesrepublik 1988 bei etwa 39 % lag. Die Unterschiede zwischen den Werten bei Projektleitern und -bearbeitern weisen darauf hin, daß auch in der Sozialforschung gilt, daß der Anteil von Frauen in leitender Position besonders gering ist. Immerhin liegt der Anteil der Projektleiterinnen etwa doppelt so hoch wie der Anteil der Professorinnen unter den Hochschullehrern, der 1987 knapp über 5 % lag. Auch der Anteil der Projektbearbeiterinnen ist höher als der Anteil der Frauen im akademischen Mittelbau, der 1987 20 % umfaßte.
Interessant ist auch die Frage, ob das Rekrutierungsverhalten von Projektleiterinnen sich von dem ihrer männlichen Kollegen unterscheidet. Dazu wurde zwischen Projekten mit mindestens einer Projektleiterin und Projekten mit rein männlicher Führung unterschieden und für diese beiden Gruppen der Anteil der männlichen und weiblichen Projektbearbeiter berechnet. Das Ergebnis zeigt einen überraschend starker Zusammenhang zwischen dem Geschlecht der

Projektleiter und dem der Projektbearbeiter: in Projekten mit weiblicher Leitung ist der Anteil der weiblichen Projektbearbeiter um über 30 Prozentpunkte höher als in Projekten mit ausschließlich männlicher Leitung.

Die "Empirische Sozialforschung 1989" ist im Campus-Verlag erschienen und kann im Buchhandel zum Preis von DM 75.- bezogen werden. **Forscher,** die 1889 Projekte für die Dokumentation gemeldet haben, können die Publikation über das Zentralarchiv zum Preis von DM 35.- erhalten.

Zur Zeit werden die Daten für die Dokumentation 1990 erhoben. In diesem Jahr werden die empirischen Projekte erstmalig in neuer Form und differenzierter erfaßt. Wir bitten um eine Unterstützung dieser Erhebung.

Data Protection and Data Access
Paul de Guchteneire and Ekkehard Mochmann (eds)
SWmOC/IFDO, North Holland, Amsterdam 1990 (ISBN 0-444-85714-1)

In the first part of this publication ten country reports are published on data protection and data access in the social sicences. The reports are from the USA, Canada, Great Britain, Norway, Sweden, Denmark, the Netherlands, the Federal Republic of Germany, France and Hungary. They were presented at the data protection Symposium of the Cologne Computer Conference 1988, which was a Joint Conference of "The International Conference on Data Bases in the Humanities and Social Sciences (ICDBHSS/88)", "The 3rd Annual Conference of the Association for History and Computing (AHC)" and "The 10th International Conference of the International Federation of Data Organizations for the Social Sciences (IFDO)". The authors wrote the reports according to a scheme suggested by the program Organizers Paul ***de Guchteneire*** and Ekkehard ***Mochmann.***

The second part of the publication consists of an international bibliography on data protection and data access in the social sciences. This bibliography was compiled in a project on data access sponsored by the Dutch Ministry of Science.

The International Federation of Data Organizations for the Social Sciences (IFDO) had sponsored the first international Conference to assess the impact of data protection on social research in ten countries in 1978 already. The proceedings of this Conference were published as "Data Protection and Social Science Research", edited by Ekkehard ***Mochmann*** and Paul J. ***Müller***, Campus Verlag, Frankfurt 1979. This volume provides a baseline for comparison, what changes have taken place in the past decade.

Die Ost-Stichprobe des SOEP
Konzept und Durchführung der "Basiserhebung 1990" in der DDR

von Jürgen Schupp und Gert Wagner

1. Die DDR-Stichprobe des SOEP

Im Juni 1990 wurde die "Basiserhebung 1990 in der DDR" durchgeführt, die als erste Welle der Ost-Stichprobe des Soziooekonomischen Panels SOEP (SOEP-Ost) dienen soll (vgl. zum SOEP Projektgruppe Panel 1990). Inhalte und Methoden der Ost-Stichprobe entsprechen in den Kernpunkten weitgehend denen des alten BRD-Panels, damit mit einem deutsch-deutschen Datenbestand vergleichende Analysen möglich sind. Längerfristiges Ziel des SOEP ist die Bereitstellung einer einheitlichen gesamtdeutschen Längsschnittmikrodatenbasis.

1.1. Design der DDR-Stichprobe

Die Grundgesamtheit der DDR-Basisbefragung des SOEP enthält, wie in der alten Bundesrepublik, Privathaushalte mit Haushaltsvorständen von Staatsbürgern der DDR; Ausländer werden nur in solchen Haushalten in die Untersuchung einbezogen, in denen sie mit einem DDR-Bürger zusammenwohnen.

Es handelt sich bei der DDR-Stichprobe ebenso wie in der BRD um eine Haushaltsstichprobe. Diese Haushaltsstichprobe wurde vom Umfrageinstitut[1] direkt mit Hilfe des bewährten Random-route-Verfahrens zur Auswahl von Haushalten realisiert. Im Gegensatz zur Bundesrepublik wurden als Startadressen solche Haushalte genommen, die in einer personenbezogenen Zufalls-Adressen-Stichprobe enthalten sind, die aus der zentralen Einwohnermeldekartei der DDR zur Verfügung gestellt wurde.

1.2. Themen- und Befragungsschwerpunkte

Das Fragenprogramm der DDR-Basisbefragung entspricht im Grundsatz dem des alten SOEP. Adaptionen der Fragen und Antwortvorgaben an spezielle DDR-Gegebenheiten sind selbstverständlich, um nicht Gefahr zu laufen, mit "systemfremden" Erhebungsinstrumenten auf Ablehnung zu stoßen. Da in der DDR im Bereich der Werte und Präferenzen besondere Veränderungen zu erwarten sind, wurde der Bereich subjektiver Indikatoren gegenüber dem

1 Die Stichprobenbildung, Durchführung der Befragung sowie die methodische Dokumentation liegen in der Verantwortung von Infratest Sozialforschung, München, das seit dem 01.07.1990 mit "Infratest Burke Berlin" zusammenarbeitet.

bundesdeutschen SOEP-Konzept (vorerst) verstärkt in das Erhebungsprogramm aufgenommen. Pragmatisch bot sich eine starke Anlehnung an die Wohlfahrts-Surveys an. Zudem sollten im Themenbereich Arbeitsmarkt und Beschäftigung die schattenwirtschaftlichen Aktivitäten in der DDR vor und nach der Währungsunion erfaßt werden.

Eine Besonderheit des SOEP besteht darin, daß nicht nur Fragen gestellt werden, die sich auf den Erhebungszeitpunkt beziehen, sondern es werden auch "retrospektive Fragen" gestellt, die sich auf das letzte Kalenderjahr beziehen. Zentral sind die sog. Kalendarien für Bildungs- und Erwerbsaktivitäten und Einkommen. In der Basisbefragung '90 für die DDR wurde gegenüber dem bundesdeutschen SOEP die Recall-Periode leicht verändert, indem im Befragungsmonat Juni 1990 monatsbezogene Aktivitäten für die *letzten* 12 Monate erfragt wurden (seit Juli 1989). Damit wird es bereits mit der Verfügbarkeit der Querschnittsdaten der ersten Welle möglich sein, die Zeit der entscheidenden gesellschaftlichen Umbrüche in der DDR im Jahre 1989/90 analytisch abzudecken. Diese Technik der retrospektiven Befragung wird es aber auch erlauben, die Zeit kurz nach der Währungsunion mit dem um dann weitere 12 Monate fortgeführten Kalendarium in der zweiten Welle zu erfassen.

Das Schul-, Ausbildungs- und Berufswesen der DDR unterscheidet sich nicht nur in der Vergangenheit, sondern auch gegenwärtig und wahrscheinlich noch für eine ganze Weile vom alten bundesdeutschen System. Entsprechend müssen institutionelle Besonderheiten für eine Befragung berücksichtigt werden, zugleich aber auch vergleichbare Klassifikationssysteme und Erhebungsinstrumente entwickelt werden. Da im Gegensatz zur Bundesrepublik für die DDR keine etablierten Indikatoren für bevölkerungsrepräsentative Surveys im Frühjahr 1990 vorlagen, mußten für die Basisbefragung '90 entsprechende - mit der BRD-Sozialstruktur vergleichbare - Indikatoren erst neu entwickelt werden. Die Lösungsmöglichkeiten, die in den Bereichen Schulabschluß, Ausbildungsabschluß und berufliche Stellung für das DDR-Panel gewählt wurden, werden bei ***Schupp und Wagner*** (1990) näher erläutert.

Ebenso wie beim SOEP in der Bundesrepublik werden die **berufliche Tätigkeit und der Wirtschaftszweig** im Klartext erhoben. Die Codierung wird von ZUMA nach der internationalen Standardklassifikation der Berufe (ISCO) sowie der SOEP-Branchenliste vorgenommen, nachdem dort erste Erfahrungen mit den Angaben der früheren Tätigkeit von Übersiedlern für ein Projekt des innerdeutschen Ministeriums gemacht wurden. Die ZUMA-Codierung von Klarschriftangaben hat den großen Vorzug, daß nach Vorliegen von Erfahrungen für den Vergleich der DDR und der BRD bzw. für internationale Vergleiche für diese Klartextangaben ggf. eine nachträgliche Recodierung mit anderen Klassifikationsschlüsseln vorgenommen werden kann.

Ein besonderes Schwergewicht bei sozialwissenschaftlichen Befragungen muß in Ostdeutschland im Bereich der **Schattenwirtschaft und Nebenerwerbstätigkeit** liegen. Dieser Bereich spielte in der DDR nicht nur in der Vergangenheit eine große Rolle. Im Zuge einer starken Wirtschafts- und Arbeitsmarktdynamik ist zu erwarten, daß die Schattenwirtschaft noch mittelfristig in Ostdeutschland eine große Bedeutung haben wird, da schattenwirtschaftliche Aktivitäten und Zweitjobs mit dazu beitragen, den Aufholprozeß der Privathaushalte der ehemaligen DDR gegenüber der alten Bundesrepublik Deutschland zu beschleunigen. Für die DDR-Basisbefragung wurden die entsprechenden Fragen nach "Eigenarbeit" und "Schattenwirtschaft" noch um den Bereich der "Beschaffungswirtschaft" erweitert.

2. Erste Ergebnisse der Feldarbeit der Basiserhebung 1990

Die Feldarbeit der Basiserhebung konnte wie geplant nahezu vollständig im Juni 1990 abgeschlossen werden. Insgesamt liegen 2179 auswertbare Haushaltsinterviews mit 4453 befragten Personen vor. Von diesen Interviews konnte 38 v.H. in der ersten und 57 v.H. in der zweiten Junihälfte durchgeführt werden. Nur in 3 v.H. der Haushalte konnte die Befragung erst Anfang Juli abgeschlossen werden. Diese Interviews, die nach Inkrafttreten des ersten Staatsvertrages zwischen der BRD und der DDR durchgeführt wurden, stören das Erhebungskonzept nicht, da der mögüche Einfluß der Währungsunion auf die subjektiven Indikatoren mit Hilfe des Interview-Datums, welches im Analyse-Datensatz enthalten sein wird, kontrolliert werden kann. Umgekehrt wäre es falsch gewesen, Haushalte nicht mehr in die Befragung einzubeziehen, die nicht bis zum 30.6. vollständig geantwortet hatten, da dadurch Gruppen, die schwierig zu interviewen sind, in der Stichprobe unterrepräsentiert worden wären.

Die Ausschöpfung der eingesetzten Adressen ist in der DDR ungewöhnlich hoch: 70 v.H. der angesprochenen Haushalte waren bereit mit allen 16jährigen und älteren Haushaltsmitgliedern an der Befragung teilzunehmen. Zum Vergleich: im Jahre 1984 waren bei der ersten SOEP-Welle in der Bundesrepublik - trotz einer weit längeren Feldzeit - nur 60 v.H. der Haushalte mit deutschem Haushaltsvorstand bereit, bei der Befragung mitzumachen.
Ein erster Vergleich der Strukturen der Stichprobe mit den Daten der amtlichen Statistik der DDR zeigt keine bedeutsamen Abweichungen. Die Teilnahmebereitschaft war - erwartungsgemäß - in Berlin geringer als in der restlichen DDR; außerdem in Gemeinden mit 2000 bis 9999 Einwohnern. Ebenso wie in der Bundesrepublik sind ältere Personen ab 60 Jahren, insbesondere über 70jährige unterrepräsentiert. Für Querschnittsauswertungen ist eine entsprechende Umgewichtung der Stichprobendaten notwendig.

Die anonymisierten Mikro-Daten der DDR-Basisbefragung werden ab Januar 1991 zur Verfügung stehen. Aus datenschutzrechtlichen Gründen ist der Abschluß eines Datenweitergabever-

träges notwendig. Anfragen sind an die Panel-Gruppe im Deutschen Institut für Wirtschaftsforschung (DIW), Königin Luise Str. 5, W-1000 Berlin 33, Tel. 030/82991283 zu richten.

Literatur

Projektgruppe 'Das Sozio-ökonomische Panel' 1990:
Das Sozio-ökonomische Panel nach fünf Wellen,
in: Vierteljahreshefte zur Wirtschaftsforschung, Heft 2, S. 141-151.

Schupp, Jürgen und *Wagner*, Gert 1990:
Die DDR-Stichprobe des Sozio-ökonomischen Panels - Konzept und Feldarbeit der Basiserhebung 1990, DIW Diskussionspapier Nr. 15, Berlin

Jürgen Schupp und Gert Wagner
Deutsches Institut für Wirtschaftsforschung
Königin Luise Str. 5
W-1000 Berlin 33.

Drogenkonsum und Drogenpolitik: Europäische Zusammenarbeit im Bereich der sozialwissenschaftlichen Drogenforschung

Organisiert von der Sektion "Soziale Probleme und soziale Kontrolle" der Deutschen Gesellschaft für Soziologie, in Kooperation mit dem Zentralarchiv für empirische Sozialforschung der Universität zu Köln, fand unter der Leitung von Karl-Heinz ***Reuband*** vom 27. bis 29. September 1990 eine Tagung europäischer und amerikanischer Wissenschaftler zum Thema "Drug Use and Drug Policy. A European Perspective" statt. Die Tagung war die erste größere internationale Veranstaltung in Europa, die einer sozialwissenschaftlichen Perspektive folgte und Fragen der Grundlagenforschung in vergleichender Perspektive gewidmet war. Referenten und Teilnehmer kamen aus der Bundesrepublik einschließlich der ehemaligen DDR, den Niederlanden, Dänemark, Schweden, Norwegen, Finnland, Großbritannien, Italien, Frankreich, Ungarn sowie den USA. In insgesamt 36 Vorträgen wurden mehrere Themenkomplexe behandelt: Epidemiologie und Ätiologie des Drogenkonsums, Drogen und Kriminalität, Aids und Needle Sharing, Drogenmortalität sowie Drogenpolitik.

Die international vergleichende Perspektive wurde einerseits ermöglicht durch Beiträge aus verschiedenen Ländern zur gleichen Thematik, andererseits durch Beiträge, die von vornherein einen Mehrländervergleich zur Basis hatten. In diesem Zusammenhang wurde u.a. auf die unterschiedliche Formen der Drogengesetzgebung in West-Europa eingegangen sowie auf die Bedeutung der wahrgenommenen Drogenrisiken für die Drogenprävalenz in der Bundesrepublik, Dänemark und den USA. Darüberhinaus wurde versucht, allgemeine Bezüge zur Entwicklung sozialer Probleme und ihrer Problemdefinition - insbesondere zu anderen Formen der Kriminalität - herzustellen.

Dem allgemeinen Bedürfnis nach verstärkter Kooperation von Sozialwissenschaftlern im europäischen Kontext folgend, kam es im Anschluß an die Tagung zur Gründung einer Arbeitsgruppe. Sie hat das Ziel, Kontakte zwischen den Sozialwissenschaftlern im Rahmen von Tagungen und gemeinsamen Forschungsprojekten zu fördern. Damit sollen die Voraussetzungen für ein verstärktes europäisches Zusammenarbeiten in der sozialwissenschaftlichen Drogenforschung geschaffen werden. Zugleich soll die in vielen Ländern bestehende Isolierung sozialwissenschaftlicher Drogenforschung durch die Ausbildung einer "scientific community" auf europäischer Ebene reduziert werden.

Der Mitgliederkreis umfaßt Soziologen, Sozialpsychologen, Historiker und Kriminologen. Neben Karl-Heinz ***Reuband*** als Vorsitzendem gehören dem Vorstand an:
Luigi ***Solivetti*** (Italien), Philip ***Bean*** (Großbritannien), Ragnar ***Hauge*** (Norwegen) und Dirk ***Korf*** (Niederlande).

Nähere Informationen bei Priv. Doz. Dr. Karl-Heinz ***Reuband,*** Zentralarchiv für empirische Sozialforschung, Universität zu Köln, Bachemer Str. 40, 5000 Köln 41, Tel. 0221-4703155.

7th European Meeting of the Psychometric Society in Trier from 29. - 31. Juli 1991

Aims **of the Conference**

The 7th European Meeting of the Psychometric Society should facilitate worldwide exchange of ideas between Psychometricians as well as Psychologists and Psychometricians. A special effort will be made to involve scientists from Eastern and Southern Europe in the Conference. Everyone reading this announcement is asked to send a copy of this 'Call for Papers' to colleagues they know and who might be interested. Special support will be offered to Eastern Europeans upon their request.

Location

The 7th European Meeting of the Psychometric Society will be held at the University of Trier, Germany, July 29-31, 1991. The Conference will take place in Building A/B of the university.

Fees

The registration fee is DM 200 (about US $ 125) before May 1, 1991. This includes a book of abstracts and free coffee, tea, and other drinks served during the breaks. A fee of 250 DM (about US $ 157) applies, if payments are received after May 1, 1991. A fee of DM 50 (US $ 31) will be charged for cancellations. Special support is offered for Eastern Europeans upon their request. Informations for modes of payment will be given in the Second Announcement.

Call for papers

Submissions should be related to one of the following areas: Test Theory, Data Analysis, Multidimensional Scaling, Statistical Methods, Simultaneous Equation Models, Correspondence Analysis, Measurement Theory, Multivariate Analysis, Mathematical Models, Factor Analysis, Psychophysical Scaling, Classification.

March 1,1991 is the deadline

for receipt of paper and Symposium proposals. Submissions of paper and Symposium proposals as well as requests for information should be sent to:

7th European Meeting
Rolf Steyer
University of Trier, FB I - Psychology
D-5500 Trier
Phone: 0651/2012056, Bitnet tfbll3 at dkluni0l

Integrating the European Data Base:
A Planning Session in the ECPR Joint Sessions of Workshops at the University of Essex, Colchester, England, 22nd - 28th March 1991

For over twenty years national data archives have collaborated to build accessible national data resources and to coordinate international data projects. The "single European market" heralds a demand to integrate the nation-specific data flies to provide a Consolidated resource to support quantitative research on a unified Europe.

This Planning Session will lay the groundwork for a Workshop at the 1992 ECPR meeting on the potential offered by existing national data resources for research on European-wide Problems. Although sponsored by the Committee of European Social Science Data Archives (CESSDA), participation is welcomed from anyone interested in developing coherent resources for quantitative comparative trans-European social research.

The Workshop scheduled for 1992 will seek contributions (a) on comparative methodology for constructing homogeneous cross-national datafiles, (b) about existing cross-national files and (c) on the priorities for establishing harmonised cross-national data resources. The Planning Session will elaborate on these topics by identifying the fields from which contributions should be sought.

Potential participants are invited to contact the Planning session convenors by December 29, 1990 at the latest. The deadline for the circulation of papers is Ist March 1991.

Convenors:

Ekkehard ***Mochmann,***
Zentralarchiv für
empirische Sozialforschung
Bachemer Str. 40
5000 Koeln 41
West Germany

Eric ***Tanenbaum***
ESRC Data Archive
University of Essex
Wivenhoe Park
Colchester, Essex C04 3SQ
England

Auf dem Weg zum Europäischen Sozialwissenschaften-Netz (ESONET)

Die sozialwissenschaftlichen Datenarchive nutzen seit Jahren Computernetze für den nationalen und internationalen Datentransfer. Fach- oder problemspezifisch organisierte Computernetze schaffen mit online-Unterstützung die kostengünstigen und technisch realisierbaren Voraussetzungen für Daten-, Informations- und Wissensvermittlung [1].

Wenn Innovation und gesellschaftlicher Fortschritt auch abhängig sind von der effizienten Kommunikation von Ideen, dann wird die Bedeutung des Informations- und Kommunikationssystems als zentralem Nervensystem moderner Gesellschaften deutlich. Modernes Kommunikationsverhalten wird aber nicht alleine deshalb zukunftsweisend, weil es die Einbahnstraße aufhebt und nun auch Signale in die andere Richtung laufen. Vielmehr müssen diese auch wahrgenommen und verstanden werden. Zukunftsweisende Kommunikation verlangt dafür offene Systeme, nicht Systeme, die Informationsproduzenten und -konsumenten voneinander abschotten. Ohne kommunikative Verschränkung entwickeln sich Informationsproduzenten und -konsumenten auseinander, Synergie findet nicht statt. Anstelle der Einbahnstraße zum zentralen Host müssen dezentrale, aber vernetzte Strukturen treten, die nicht nur Informationsabruf, sondern auch Kommunikation ermöglichen [2].

Die technische Entwicklung bietet heute die nötigen Grundlagen.
PCs werden immer breiter verfügbar. Mit der Digitalisierung des Fernsehens ist auch hier der Schritt zum "intelligenten Chip" nicht mehr weit. Das Telefon steht ohnehin in den meisten Haushalten daneben. Für Abteilungen oder Kleinbetriebe stehen leistungsfähige Microrechner oder auch Workstations zur Verfügung. Großunternehmen und Universitäten verfügen nach wie vor über Superrechner. Die moderne Telekommunikation macht nun Länder- und Erdteile übergreifend die Vernetzung aller Komponenten möglich und diese schreitet rapide voran. Praktisch kann jeder Teilnehmer am Informationsaustausch mit jedem anderen gleichberechtigt verbunden werden.

Das Wissenschaftsnetz WIN des DFN-Vereins (Deutsches Forschungsnetz) stellt zusammen mit dem europäischen IXI-Netz ein attraktives Angebot zur Lösung der Datentransfer- und Kommunikationsaufgaben auch im sozialwissenschaftlichen Bereich. Um sowohl den einzelnen Wissenschaftler als auch Projekte mehrerer Institute optimal mit Hilfe neuester Telekommunikationstechniken unterstützen zu können, haben sich mit den GESIS-Instituten weitere Institute zur Gründung einer DFN-Nutzergruppe "Sozialwissenschaften/Empirische Sozialforschung" abgestimmt. Unter Federführung des Zentralarchivs wurde ein Antrag an das Deutsche Forschungsnetz zur Schaffung eines Sozialwissenschaftennetzes (SONET) erarbeitet.
Im Rahmen eines zweijährigen Projektes sollen die Voraussetzungen für die moderne Tele-

kommunikation im Bereich Daten- und Informationstransfer, Vernetzung von Datenbanken sowie die Unterstützung der Kommunikation zwischen den Wissenschaftlern geschaffen werden.

Vom Committee of European Social Science Data Archives (CESSDA) wurde im September in Köln vorgeschlagen, die Möglichkeiten der Wissenschaftsnetze auch im internationalen Verbund weiter zu entwickeln. Das Wissenschaftsnetz (WIN) des DFN stellt zusammen mit dem europäischen DCI-Netz hierfür die technischen Basisdienste zur Verfügung. Damit bietet sich die Chance zur Entwicklung eines europäischen Sozialwissenschaftennetzes (ESONET).

Anmerkungen:

(1) ***Mochmann,*** Ekkehard; ***Rau,*** Friedrich
Rechnerverbundnetze für den sozialwissenschaftlichen Datenservice.
In: ZA-Information, 1984, Nr. 14, S. 59-61

Mochmann, Ekkehard; ***Rose***, Erwin; ***Uher***, Rolf
Computernetze für die Sozialforschung.
In: ZA-Information, 1985, Nr. 16, S. 95-98

Bauske, Franz
DFN-EARN-Verbund - Zentralarchiv als Gateway zwischen zwei Computernetzen
In: ZA-Information, 1986, Nr. 18, S. 20 - 25

Schnell, Rainer
Möglichkeiten der Nutzung von BITNET in den Sozialwissenschaften.
In: ZA-Information, 1989, Nr. 24, S. 101-115

(2) ***Mochmann,*** Ekkehard
Moderne Netzwerke zur nationalen und internationalen Datenkommunikation
In: Informationsmethoden: Neue Ansätze und Techniken
Deutscher Dokumentartag 1989 - Proceedings, Deutsche Gesellschaft für Dokumentation, Frankfurt 1990, S. 117-128

Ekkehard Mochmann, Friedrich Rau

Multivariate Analyse kategorialer Daten

Frühjahrsseminar des Zentralarchivs für empirische Sozialforschung
vom 18. Februar bis 8. März 1991

Das Frühjahrsseminar ist ein Fortbildungsangebot für Sozialwissenschaftler, die Kenntnisse in fortgeschrittenen Techniken der Datenanalyse erwerben und in der Anwendung an Daten erproben wollen. Statistische Grundkenntnisse und Erfahrungen mit statistischen Auswertungsprogrammen werden vorausgesetzt. Das Seminar besteht aus Vorlesungen und Übungen. In den Vorlesungen werden die Logik von Modellen und die darauf aufbauenden Analyseverfahren erläutert. In den Übungen sollen die Teilnehmer die in der Vorlesung dargestellten Verfahren praktisch anwenden. Während des Frühjahrsseminars besteht darüber hinaus die Gelegenheit, die Funktionen und Serviceangebote des Zentralarchivs kennenzulernen.

Thema des Frühjahrsseminars 1991 ist die Analyse nicht-metrischer Daten mit Hilfe log-linearer und logistischer Modelle. Damit soll einem Datentyp Rechnung getragen werden, der in der Sozialforschung oft vorkommt, der aber mit Hilfe der bisher verbreiteten Analyseverfahren in vielen Fällen nicht angemessen analysiert werden kann. Im Verlauf des Frühjahrsseminars werden drei Verfahren behandelt:

- **die log-lineare und logistische Analyse**

- **die logistische Analyse mit Individualdaten**

- **die Analyse von Längsschnittdaten mit Hilfe latenter Variablen**

Die erste Woche (log-lineare und logistische Modelle) besteht aus einer Einführung in die gesamte Modellklasse. Ausgehend von einer konkreten Fragestellung, in der alle untersuchten Merkmale nicht-metrisches Skalenniveau haben, wird die häufigste Anwendung - die Tabellenanalyse - besprochen.

Bei der Analyse von Kreuztabellen wird vorausgesetzt, daß alle untersuchten Merkmale nominales oder ordinales Skalenniveau haben - metrische Variablen müssen ggfs. klassifiziert werden. Diese Datenkonstellation trifft bei sozialwissenschaftlichen Untersuchungen zwar häufig zu, jedoch möchte man in einigen Fällen metrische und nicht-metrische Merkmale gemeinsam betrachten.

Gegenstand der zweiten Woche (logistische Modelle für Individualdaten) sind daher Modelle für nicht-metrische Zielvariablen (z.B. Parteipräferenz), die sowohl durch andere nicht-metrische (z. B. Bildungsabschluß) als auch durch metrische Merkmale (z. B. Einkommen) vorhergesagt werden sollen. Dieser Modelltyp ist auch unter der Bezeichnung "Modelle diskreter Wahl" bekannt.

In der dritten Woche (Analyse von Längsschnittdaten) wird es vor allem um die Anwendung log-linearer Modelle bei der Analyse von Längsschnittdaten gehen. Die log-lineare Analyse wird hier durch die Einführung latenter Variablen erweitert. Ein besonderes Problem von Längsschnittdaten ist, daß sozialwissenschaftliche Messungen häufig ungenau sind, so daß Änderungen auftreten, die allein auf die Erhebungsmethode zurückgehen. Um tatsächliche von solchen scheinbaren Veränderungen trennen zu können, ist es daher notwendig, die Zuverlässigkeit der Daten zu kontrollieren. Ein weiteres Thema dieses Vorlesungsteils wird der Umgang mit der Panelmortalität sein.

Personen, die schon gewisse Vorkenntnisse haben, können (je nach Interessenschwerpunkt) auch nur an der zweiten oder dritten Woche teilnehmen. Allen anderen wird jedoch unbedingt empfohlen, die Einführung in der ersten Woche zu besuchen. In den Vorlesungen werden u. a. die statistischen Annahmen der Verfahren und die Konsequenzen bei der Verletzung der Annahmen diskutiert. Programme für Auswertungen mit log-linearen und/oder logistischen Modellen findet man in allen großen sozialwissenschaftlichen Analysepaketen. Für die dritte Veranstaltungswoche ist ein stand-alone-Programm (LCAG) vorgesehen. Die programmtechnische Umsetzung wird im Vorlesungsteil behandelt und in den Arbeitsgruppen an ausgewählten Datensätzen eingeübt. Am Sonntagnachmittag vor Beginn des jeweils neuen Vorlesungsteils wird zur Auffrischung von Grundkenntnissen im Bereich der computerunterstützten Datenanalyse ein Workshop angeboten.

Die Verfahren werden im Rahmen getrennter Vorlesungen vorgestellt. Die Vorlesungen werden gehalten von:

18. - 22. Februar 1991:
Log-lineare und logistische Analyse
Prof. Dr. H. J. ***Andress*** (Universität Bielefeld)

25. Februar-1. März 1991:
Logistische Analyse mit Individualdaten
Prof. **Dr. H. J. *Andress*** (Universität Bielefeld)

4. März-8. März 1991:
Analyse von Längsschnittdaten mit Hilfe latenter Variablen
Prof. **Dr. J. *Hagenaars*** (Universität Tilburg)

Zusätzlich werden Gastreferate zu inhaltlichen und methodischen Problemen sowie über die Dienstleistungen des Zentralarchivs geboten. Die praktische Umsetzung im Rahmen von Arbeitsgruppen wird unterstützt von Jörg ***Blasius,*** Steffen ***Kühnel*** und Karl-Heinz ***Reuband*** aus dem Zentralarchiv. Die drei Arbeitsgruppen verwenden die Wahlstudie 1987 als Datenbasis. Sofern bereits verfügbar, werden darüber hinaus die Daten der neuesten Bundestagswahlstudie 1990 benutzt. Dabei stehen jeweils ausgewählte Variablen zu unterschiedlichen Themen zur Verfügung: zu "cleavage"-Strukturen, Protestbewegungen, allgemeinen sozialen und politischen Einstellungen. Je nach Arbeitsgruppe werden stärker inhaltliche oder methodische Fragen im Vordergrund stehen. Die Arbeitsgruppenzuordnung wird jeweils zu Beginn der Woche stattfinden.

Für die Teilnahme über die gesamten drei Wochen wird eine Gebühr von 200,- DM erhoben. Bei Teilnahme nur an Teilen des Frühjahrsseminars beträgt der Kostenbeitrag DM 75,-- pro Woche. In Ausnahmefällen ist eine Ermäßigung möglich. Fahrt- und Aufenthaltskosten sind von den Teilnehmern zu tragen. Benutzen Sie für die Anmeldung das Formular auf den folgenden Seiten und senden Sie es bitte bis **spätestens** 10. Januar 1991 an das Zentralarchiv. Da die Teilnehmerzahl begrenzt ist, wird eine möglichst umgehende Anmeldung empfohlen.

Bitte abtrennen!

An
Zentralarchiv
für empirische Sozialforschung
Universität zu Köln
Bachemer Str. 40
5000 Köln 41

Anmeldung zum Frühjahrsseminar 1991

Name _____

Anschrift _____

_____

_____

Telefon _____

Telefax _____

Besuchte Universität _____

Studienabschluß _____

Derzeitige Tätigkeit _____

_____

Institution _____

_____

Aufgabenbereich ❒ Lehre ❒ Forschung

Mit welchen Datenanalyseprogrammen und Programmpaketen sind Sie näher vertraut?

Für Lehrende:
Welche Veranstaltungen haben Sie in den letzten Jahren durchgeführt (Inhalt, Typ)?

Teilnahme

❒ am gesamten Frühjahrsseminar vom 18.02 bis 8.3.1991

nur an Teilen des Frühjahrsseminars (bitte Zutreffendes ankreuzen)

❒ ***Andress:*** Log-lineare und logistische Analyse (18. - 22.2.91)

❒ ***Andress:*** Logistische Analyse mit individualdaten (25.2. -1.3.91)

❒ ***Hagenaars:*** Analyse von Längsschnittdaten mit Hilfe latenter Variablen (4.3. - 8.3.91)

Anmeldeschluß: 10. Januar 1991

Datum _____ Unterschrift _____